

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12Q 1/68</b>	<b>A1</b>	(11) International Publication Number: <b>WO 99/60164</b> (43) International Publication Date: 25 November 1999 (25.11.99)
<p>(21) International Application Number: PCT/US99/11066</p> <p>(22) International Filing Date: 14 May 1999 (14.05.99)</p> <p>(30) Priority Data: 60/085,673 15 May 1998 (15.05.98) US</p> <p>(71) Applicant: QUARK BIOTECH, INC. [US/US]; 1059 Serpentine Lane, Pleasanton, CA 94566 (US).</p> <p>(72) Inventors: EINAT, Paz; Neve Nir 1/27, Ness-Ziona (IL). MOR, Omer; Emek Ayalon Street 12, Ganei Ilan, Kiryat Ono (IL). SKALITER, Rami; Habaniam Street 117/10, Ness-Ziona (IL). FEINSTEIN, Elena; Hahagana Street 12/29, Rehovot (IL). FAERMAN, Alexander; Haeliah 2024/36, Bnei Aish (IL).</p> <p>(74) Agents: KOWALSKI, Thomas, J. et al.; Frommer Lawrence &amp; Haug LLP, 745 Fifth Avenue, New York, NY 10151 (US).</p>	<p>(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p>	
<p><b>Published</b> With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</p>		
<p>(54) Title: MECHANICAL STRESS INDUCED GENES, EXPRESSION PRODUCTS THEREFROM, AND USES THEREOF</p>		
<p>(57) Abstract</p>		
<p>This disclosure relates to osteoporosis. Moreover, the disclosure relates to mechanical stress induced genes, probes therefor, tests to identify such genes, expression products of such genes, uses for such genes and expression products, e.g., in diagnosis (for instance risk determination), treatment, prevention, or control, of osteoporosis or factors or processes which lead to osteoporosis; and, to diagnostic, treatment, prevention, or control methods or processes, as well as compositions therefor and methods or processes for making and using such compositions.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NZ	New Zealand	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroun	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**TITLE OF THE INVENTION**

***MECHANICAL STRESS INDUCED GENES,  
EXPRESSION PRODUCTS THEREFROM,  
AND USES THEREOF***

**5 RELATED APPLICATION**

This application is based upon and claims priority from U.S. Provisional application Serial No. 60/085,673, filed May 15, 1998.

Reference is also made to U.S. Provisional application Serial No. 60/084,944, filed May 11, 1998, and the full U.S. utility application, Serial No. \_\_\_\_\_, filed  
10 May 11, 1999, naming as inventors Paz Einat, Rami Skaliter, Orna Mor and Sylvie Luria and assigned to the assignee of the present application (Kohn & Associates Attorney Docket No. 0168.00060), and claiming priority from U.S. Provisional application Serial No. 60/084,944 (herein "the May 11, 1999 Einat et al. full U.S. utility application").

15 U.S. Provisional application Serial No. 60/085,673, filed May 15, 1998, U.S. application Serial No. 60/084,944, and the May 11, 1999 Einat et al. full U.S. utility application, as well as each document or reference cited in that application, is hereby expressly incorporated herein by reference. Documents or references are also cited in the following text, either in a Reference List before the claims, or in the text itself;  
20 and, each of these documents or references ("herein-cited documents or references"), as well as each document or reference cited in each of the herein-cited documents or references, is hereby expressly incorporated herein by reference. It is explicitly stated that the inventive entity of the May 11, 1999 Einat et al. full U.S. utility application is not another or others as to the inventive entity of this application; and, that the  
25 inventive entity of the present application is not another or others as to the inventive entity of the May 11, 1999 Einat et al. full U.S. utility application.

**FIELD OF THE INVENTION**

This invention relates to osteoporosis. Moreover, the invention relates to mechanical stress induced genes, probes therefor, tests to identify such genes,  
30 expression products of such genes, uses for such genes and expression products, e.g., in diagnosis (for instance risk determination), treatment, prevention, or control, of osteoporosis or factors or processes which lead to osteoporosis; and, to diagnostic,

treatment, prevention, or control methods or processes, as well as compositions therefor and methods or processes for making and using such compositions.

The present application also relates to a method for identifying genes that are regulated at the RNA level. More specifically, the present application relates to the rapid isolation of differentially expressed or developmentally regulated gene sequences through analysis of mRNAs obtained from specific cellular compartments. By comparing changes in the relative abundance of the mRNAs found in these compartments occurring as a result of application of a cue or stimulus to the tested biological sample, genes that are differentially expressed can be characterized.

The present invention especially relates to such methods with respect to bone cells and/or the stimulus being mechanical stress.

These and other areas to which the invention relates will be apparent from the following text.

#### **BACKGROUND OF THE INVENTION**

Bone is composed of a collagen-rich organic matrix impregnated with mineral, largely calcium and phosphate. Two major forms of bone exist, compact cortical bone forms the external envelopes of the skeleton and trabecular or medullary bone forms plates that traverse the internal cavities of the skeleton. The responses of these two forms to metabolic influences and their susceptibility to fracture differ.

Bone undergoes continuous remodeling (turnover, renewal) throughout life. Mechanical and electrical forces, hormones and local regulatory factors influence remodeling. Bone is renewed by two opposing activities that are coupled in time and space (Parfitt 1979). These activities - resorption and formation - are contained within a temporary anatomic structure known as a bone remodeling unit (Parfitt 1981). Within a given bone remodeling unit, old bone is resorbed by osteoclasts. The resorbed cavity created by the osteoclasts is subsequently filled with new bone by osteoblasts, which synthesize the organic matrix of bone.

Peak bone mass is mainly genetically determined, though dietary factors and physical activity can have positive effects. Peak bone mass is attained at the point when skeletal growth ceases, after which time bone loss starts.

In contrast to the positive balance that occurs during growth, in osteoporosis, the resorbed cavity is not completely refilled by bone (Parfitt 1988). Osteoporosis, or porous bone, is a progressive and chronic disease characterized by low bone mass and



structural deterioration of bone tissue, leading to bone fragility and an increased susceptibility to fractures of the hip, spine, and wrist (diminishing bone strength).

Bone loss occurs without symptoms. The Consensus Development Conference (Am J Med 94:646-50, 1993) defined osteoporosis as "a systemic skeletal disease characterized by low bone mass and microarchitectural deterioration of bone tissue, with a consequent increase in bone fragility and susceptibility to fracture."

Common types of osteoporosis include postmenopausal osteoporosis; and senile osteoporosis, which generally occurs in later life, e.g., 70+ years; *see, e.g.*, U.S. Patent No. 5,691,153 and documents cited therein and during its prosecution, all incorporated herein by reference.

Osteoporosis is estimated to affect more than 25 million people in the United States (Rosen 1997); and, at least one estimate asserts that osteoporosis affects 1 in 3 women (Keen et al. 1997). However, life expectancy has increased, and in the western world, 17% of women are now over 50 years of age; and, a woman can expect to live a third (1/3) of her life after menopause. Thus, some estimate that 1 out of every 2 women and 1 out of 5 men will eventually develop osteoporosis; and, that 75 million people in the US, Japan and Europe have osteoporosis. The World Summit of Osteoporosis Societies estimates that more than 200 million people world-wide are afflicted with the disease. The actual incidence of the disease is difficult to estimate since the condition is often asymptotic until a bone fracture occurs. It is believed that there are over 1.5 million osteoporosis-associated bone fractures per year in the U.S. of which 300,000 are hip fractures that usually require hospitalization and surgery and may result in lengthy or permanent disability or even death. (*See* Spangler et al. "The Genetic Component of Osteoporosis Mini-review;

<http://www.csa.com.osteointro.html>)

Further, there is a 20-30% mortality rate related to hip fractures in elderly women (U.S. Patent No. 5,691,153); and, it is reported that such a patient with a hip fracture has a 10-15% greater chance of dying than others of the same age. Further, it is reported that although men suffer fewer hip injuries than women, men are 25% more likely than women to die within one year of the injury. *See* Sprangler et al., *supra*. Also, about 20% of the patients who were living independently before a hip fracture still remain confined in a long-term health care facility one year later; and, the treatment of osteoporosis and related fractures can cost over \$10 billion annually.

Accordingly, osteoporosis is a major health problem in virtually all societies (Eisman 1996; Wark 1996; U.S. Patent No. 5,834,200 and the documents cited therein, being hereby incorporated herein by reference).

Treatment for osteoporosis helps stop further bone loss and fractures, such as  
5 HRT (hormone replacement therapy), bisphosphonates, e.g., alendronate (Fosamax), as well as, estrogen and estrogen receptor modulators, progestin, calcitonin, and vitamin D.

While there may be numerous factors that determine whether any particular person will develop osteoporosis, a step towards prevention, control or treatment of  
10 osteoporosis is determining whether one is at risk for osteoporosis. Genetic factors are said to play an important role in the pathogenesis of osteoporosis (Ralston 1997; *see also* Keen et al. 1997; Eisman 1996; Rosen 1997; Cole 1998, Johnston et al. 1995; Gong et al. 1996; Wasnich 1996 *inter alia*).

Some attribute 50-60% of total bone variation (Bone Mineral Density; BMD),  
15 depending upon the bone area, to genetic effects (Livshits et al. 1996). However, up to 85%-90% of the variance in bone mineral density may be genetically determined.

For instance, as studies have shown from family histories, twin studies, and racial factors, there may be a predisposition for osteoporosis (*see, e.g.,* Jouanny et al. 1995; Garnero et al. 1996; Cummings 1996; Lonzer et al. 1996). Several candidate  
20 genes may be involved in this, most probably multigenic process.

Association between vitamin D receptor gene (VDR) allelic variation and BMD has been suggested. Restriction fragment length polymorphisms (RFLPs) at the vitamin D receptor (VDR) gene locus have been recently correlated to bone mineral density (BMD) and rate of bone loss (*see, e.g.,* Tokita et al. 1996; Cole et al. 1998;  
25 Eisman 1996; Keen et al. Ralston 1997; Fujita 1996; Houston et al. 1996; Riggs et al. 1995; Fleet et al. 1995; Krall et al. 1995).

Collagen type I alpha gene has been implicated (*see, e.g.,* Dalglish 1997; Pereira et al. 1995). The COL1A1 and COL1A2 genes encode type I collagen, a key bone protein, and, therefore, may play a role in the genetic control of bone mass.

30 Mutation of the estrogen receptor (ER) gene may be implicated in some cases of osteoporosis (polymorphism of the ER gene has been correlated with BMD in some populations) (*see* Sano et al. 1995; *see also* U.S. Patent No. 5,834,200). Interleukin 1 (IL-1) and tumor necrosis factor alpha (TNF-alpha) have also been

implicated in the pathogenesis of osteoporosis in recent studies. These proinflammatory cytokines induce both cyclooxygenase (COX) and nitric oxide synthase (NOS) with the release of prostaglandin (PG) and NO, respectively. Cytokines have been shown to be powerful regulators of bone resorption and formation, though under superior control from oestrogen/testosterone, parathyroidhormone and 1,25(OH)2D3. Some of the cytokines primarily enhance osteoclastic bone resorption e.g. IL-1 (Interleukin-1), TNF (Tumor Necrosis Factor) and IL-6 (Interleukin-6), while others primarily stimulate bone formation e.g. TGF-beta (Transforming Growth Factor), IGF (Insulin-like Growth Factor) and PDGF (Platelet Derived Growth Factor).

There is need for clinical and epidemiological research to further explore and extend the current potential for practical control, prevention and treatment of the disease. A deeper knowledge of factors controlling bone cell activity and regulation of bone mineral and matrix formation and remodeling is desired.

Further, while certain genetic may be useful for detecting high bone mass or predisposition to low or high bone mineral density (*see* U.S. Patents Nos. 5,691,153 and 5,834,200), there is a need for further tests to determine risk for osteoporosis; and, there is a need for new treatments, preventatives, or means to control osteoporosis or factors or processes which lead to osteoporosis.

## **OBJECTS AND SUMMARY OF THE INVENTION**

An object of the invention can include any one or any combination or all of: advancing clinical and epidemiological research and/or further exploring and extending the current potential for practical control, prevention and treatment of the disease; providing further knowledge of factors controlling bone cell activity and regulation of bone mineral and matrix formation and remodeling; providing further tests to determine risk for osteoporosis; and/or new treatments, preventatives, or means to control osteoporosis or factors or processes which lead to osteoporosis.

The present invention provides mechanical stress induced genes, probes therefor, a test to identify such genes, expression products of such genes, uses for such genes and expression products, e.g., in diagnosis (for instance risk determination), treatment, prevention, control, or osteoporosis or factors or processes which lead to osteoporosis. Thus, the invention further provides diagnostic, treatment, prevention, control methods or processes as well as compositions.

There is disclosed a method for identifying genes whose expression is regulated at the RNA level in an organism including the steps of selectively stimulating translation of an unknown target mRNA with a stress inducing element, the target mRNA being part of a larger sample of mRNA, dividing the sample of mRNA into pools of translated and untranslated mRNA and differentially analyzing the pools of mRNA to identify genes translationally regulated by the stress inducing element. The stress inducing element can include pathologic, environmental including chemical and physical stressors or other stimulus that induces mRNA translation. The stress inducing element can comprise mechanical stress. The sample can comprise bone cells that retain being bone cells in a culture, e.g., calvaria cells.

According to the present application, methods are disclosed for identifying genes that may be regulated on a number of possible regulatory levels. Such methods include the steps of exposing cells or tissue to a cue or stimulus such as mechanical, chemical, toxic, pharmaceutical or other stress, hormones, physiological disorders or disease; fractionating the cells into compartments such as polysomes, nuclei, cytoplasm and spliceosomes; extracting the mRNA from these fractions, and subjecting the mRNA to differential analysis using accepted methodologies, such as gene expression array (GEM).

For instance, the application discloses the use of RNA isolation from nuclei for isolating genes whose steady state levels show only minor changes, but which show high differential expression when detected by nuclear RNA probe. Most such genes are regulated at the transcriptional level. One type of regulation is shown using polysomes isolated from cells/tissues to identify genes whose mRNA steady state levels do not change, but are highly increased in the polysomes after application of a stress cue. Such genes are regulated strictly on the translation level. A subgroup of genes regulated on the translational level involves the existence of internal ribosome entry sites. A method is disclosed for identification of such genes, which includes inhibiting 5' cap-dependant mRNA translation in a cell, collecting a pool of mRNA from the cells, and differentially analyzing the pool of mRNA to identify genes with sequences coding for internal ribosome entry sites.

Thus, the application discloses a method or process for identifying genes whose expression is responsive to a specific cue or cues including the steps of:

(a) applying a cue to an organism or tissue or cells;

(b) isolating specific cellular fractions from the tissues or cells subjected to the cue;

(c) extracting the mRNA from the cellular fractions; and

(d) differentially analyzing the mRNA samples in comparison with control  
5 samples not subjected to the cue to identify genes that have responded to the cue.

The cells or tissues can be bone cells which retain the nature of being bone cells when in a culture and the cue can be mechanical stress or a lack thereof.

The cue can be a toxin or a chemical, or a pharmaceutical, or a mechanical stress, or an electric current, or a pathogen or a pathological condition, or a hormone,  
10 or a specific protein. The cue can be further defined as chemically treating the cells, or irradiating the cells, or depriving the cells of oxygen. The cue can be further defined as a stress-inducing element of unknown relationship to gene translation.

The genes can be identified at the translation level; genes regulated at the transcription level; genes regulated by RNA stability; genes regulated by mRNA  
15 transport rate between the nucleus and cytoplasm; genes regulated by differential splicing; and genes regulated by antisense RNA.

The mRNA samples can be further fractionated into mRNA subfractions which are subjected to differential analysis to identify genes responsive to the cue at all levels of expression regulation as herein defined and to determine the abundance  
20 and direction of the response. The mRNA sample can be fractionated into one or more subfractions from the group consisting essentially of cytoplasmic, nuclear, polyribosomal, sub polyribosomal, microsomal or rough endoplasmic reticulum, mitochondrial and spliceosome associated mRNA.

The differential analysis step can be selected from the group consisting of  
25 differential display, representational differential analysis (RDA), suppressive subtraction hybridization (SSH), serial analysis of gene expression (SAGE), gene expression microarray (GEM), nucleic acid chip technology, oligonucleotide chip technology; DNA membrane arrays; direct sequencing and variations and combinations of these methods. The differential analysis step can be further defined  
30 as identifying and measuring the genes regulated at the translation level. The differential analysis step can also be further defined as identifying and measuring the genes regulated at the transcription level. The differential analysis step can also be further defined as identifying and measuring the genes regulated by RNA stability.

The differential analysis step can additionally be further defined as identifying and measuring the genes regulated by mRNA transport rate between the nucleus and the cytoplasm. The differential analysis step can also be further defined as identifying and measuring the genes regulated by differential splicing. The differential analysis step additionally can be further defined as identifying and measuring the genes encoding secreted and membrane proteins. The differential analysis step can also be further defined as identifying and measuring the genes encoding for nuclear proteins.

The application further discloses a method for determining risk of developing a physiological or disease state based upon absence or decrease from normal cells of mRNA or protein from a gene shown to be down regulated in a mammal by an inventive or disclosed method comprising:

- (a) determining the level or status of mRNA in cells of said mammal; and/or
- (b) determining the level or status of corresponding protein in cells of said mammal; and
- (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining an absence or decrease from normal cells of mRNA or protein and thus the risk of developing a physiological or disease state.

The application further discloses a method for determining risk of physiological or disease state based upon presence or an increase from normal cells of mRNA or protein from a gene shown to be upregulated by an inventive or disclosed method in a mammal comprising:

- (a) determining the level or status of mRNA in cells of said mammal; and/or
- (b) determining the level or status of corresponding protein in cells of said mammal; and
- (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a

transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining a  
5 presence or increase from normal cells of mRNA or protein and thus risk.

These "determining" methods can be diagnostic methods; e.g., methods for diagnosing a physiological or disease state.

The application further discloses a method for testing a medicament for or a gene therapy approach to a physiological or disease state or other factors causing or  
10 contributing thereto or to symptoms thereof based on absence or decrease from normal cells or presence or increase from normal cells of mRNA or protein of identified genes comprising an inventive or disclosed method additionally comprising: (a') administering the medicament or the gene therapy; and from the comparing determining an absence or decrease from normal cells or presence or  
15 increase from normal cells of the relevant mRNA or protein and thus efficacy of the medicament or the gene therapy.

The application still further discloses a method for treating, preventing or controlling a physiological or disease state comprising an inventive or disclosed method and additionally comprising administering a medicament or treatment therefor  
20 or for a cause thereof or a symptom thereof.

The medicament or treatment can comprise the protein, a functional portion thereof, a vector expressing the protein or a functional portion thereof, or an inhibitor of the protein or of a functional portion thereof, or an inhibitor of a nucleic acid encoding the protein or a functional portion thereof.

25 Inventive or disclosed methods can further comprise:

- (d) determining the level or status of a second gene mRNA in cells of said mammal; and/or
- (e) determining the level or status of protein expressed by a second gene product in cells of said mammal; and
- 30 (f) comparing said level or status of that mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence,

which may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining risk.

- 5 Steps (a) and/or (b) and optionally (d) and/or (e) can be carried out *in vivo* and/or steps (a) and/or (b) and optionally (d) and/or (e) can be carried out *in vitro*. The determination in step (a) and optionally in step (d) can be effected by employing
- (i) a nucleic acid sequence corresponding to at least a part of the gene encoding at least part of the protein and optionally a second nucleic acid sequence corresponding
  - 10 to at least a part of the second gene encoding at least part of the second protein;
  - (ii) a nucleic acid sequence complementary to the nucleic acid sequence(s) of (i); or
  - (iii) a primer or a primer pair hybridizing to the nucleic acid sequence(s) of (i) or (ii).

- 15 The determination in step (b) and optionally of step (e) can be effected by employing an antibody or a fragment thereof that specifically binds to the protein and optionally by employing a second antibody or a fragment thereof which specifically binds to the second protein.

- In inventive methods, the stimulus can be mechanical stress or a lack thereof
- 20 and the sample comprises bone cells which retain their characteristic thereof in cultures.

- The invention further provides a gene identification process comprising: preparation of probes from a model system; analysis of DNA chip hybridization; sequencing of clones showing differential expression; and optionally full-length
- 25 cloning of clones of interest.

The model system can comprise bone cells which retain their characteristic thereof in cultures which have mechanical stress or a lack thereof applied thereto. The bone cells can comprise a calvaria primary culture.

- The invention further provides a method for determining risk of developing
- 30 osteoporosis or low or high bone density or other factors causing or contributing to osteoporosis or other conditions involving mechanical stress or a lack thereof, based upon absence or decrease from normal cells of mRNA or protein from a gene shown to be down regulated by an inventive method comprising:



- (a) determining the level or status of mRNA in bone cells of said mammal; and/or
  - (b) determining the level or status of corresponding protein in bone cells of said mammal; and
  - (c) comparing said level or status of mRNA and/or protein with the corresponding
- 5 level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including
- 10 post-translational modifications of the protein, and from the comparing determining an absence or decrease from normal cells of mRNA or protein and thus risk.

- The invention still further provides a method for determining risk of developing osteoporosis or low or high bone density or other factors causing or contributing to osteoporosis or other conditions involving mechanical stress or a lack thereof, based upon presence or increase from normal cells of mRNA or protein from
- 15 a gene shown to be upregulated by an inventive method in a mammal comprising:
- (a) determining the level or status of mRNA in bone cells of said mammal; and/or
  - (b) determining the level or status of corresponding protein in bone cells of said mammal; and
  - (c) comparing said level or status of mRNA and/or protein with the corresponding
- 20 level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity
- 25 of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining a presence or increase from normal cells of mRNA or protein and thus risk.

- The invention also provides a method for determining risk of developing osteoporosis or low or high bone density or other factors causing or contributing to
- 30 osteoporosis or lower levels of osteoblasts and chondrocytes or other conditions involving mechanical stress or a lack thereof, based upon absence or decrease from normal cells of mRNA or protein from 608 comprising:
- (a) determining the level or status of mRNA in bone cells of said mammal; and/or

(b) determining the level or status of corresponding protein in bone cells of said mammal; and

(c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining a presence or increase from normal cells of mRNA or protein and thus risk.

Further still, the invention provides a method for determining risk of developing osteoporosis or low or high bone density or other factors causing or contributing to osteoporosis or of imbalance as to osteogenic and chondrogenic cells or other conditions involving mechanical stress or a lack thereof, based upon absence or decrease from normal cells or presence or increase from normal cells of mRNA or protein from 405 in a mammal comprising:

(a) determining the level or status of mRNA in bone cells of said mammal; and/or  
(b) determining the level or status of corresponding protein in bone cells of said mammal; and

(c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining an absence or decrease from normal cells of mRNA or protein and thus risk.

Even further, the invention provides a method for determining risk of developing osteoporosis or low or high bone density or other factors causing or contributing to osteoporosis or of being susceptible to environmental factors or other than genetic factors of osteoporosis or of predisposition of bones towards susceptibility to environmental factors, or less lymphoid cells, or osteoporosis, or other conditions involving mechanical stress or a lack thereof, based upon presence or

increase from normal cells or absence or decrease from normal cells of mRNA or protein from 274 in a mammal comprising:

- (a) determining the level or status of mRNA in bone cells of said mammal; and/or
- (b) determining the level or status of corresponding protein in bone cells of said  
5 mammal; and
- (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a  
10 mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining an absence or decrease from normal cells of mRNA or protein and thus risk.

These "determining" methods can be diagnostic methods; e.g., methods for  
15 diagnosing osteoporosis or for diagnosing other conditions recited in the preamble of these "determining" methods.

Also, the invention further provides a method for testing a medicament for or gene therapy approach to osteoporosis or bone density or other factors causing or contributing to osteoporosis or symptoms thereof or other conditions involving  
20 mechanical stress or a lack thereof, based on absence or decrease from normal cells or presence or increase from normal cells of mRNA or protein comprising a method according to any one of the foregoing inventive methods and additionally comprising:  
(a') administering the medicament or the gene therapy; and from the comparing determining an absence or decrease from normal cells or presence or increase from  
25 normal cells of the relevant mRNA or protein and thus efficacy of the medicament or the gene therapy.

The invention also comprehends analogous methods with respect to other genes identified by inventive processes, e.g., CMF2-224, CMF2-45.

Similarly, the invention additionally provides a method for treating,  
30 preventing or controlling osteoporosis or other conditions involving mechanical stress or a lack thereof, comprising a method according to any one of the foregoing inventive methods and further comprising administering a medicament or treatment for osteoporosis or a cause thereof or a symptom thereof.

Still further, the invention provides a composition comprising a gene or portion thereof or a protein or portion thereof expressed by the gene or portion thereof or an antibody or portion thereof which binds to the protein or portion thereof, wherein the gene is identified by an inventive method.

- 5 Even further still, the invention provides an osteoporosis or mechanical stress or lack thereof model comprising bone cells which retain their characteristic thereof in culture with mechanical stress applied thereto or an absence of mechanical stress applied thereto.

- The invention additionally provides an isolated nucleic acid molecule  
10 encoding the herein identified protein 608 or a functional portion thereof or a polypeptide which is at least substantially homologous or identical thereto.

Further, the invention provides an isolated nucleic acid molecule encoding the herein identified protein 405 or a functional portion thereof or a polypeptide which is at least substantially homologous or identical thereto.

- 15 Also, the invention provides an isolated nucleic acid molecule encoding the herein identified protein 274 or a functional portion thereof or a polypeptide which is at least substantially homologous or identical thereto.

- The invention comprehends an isolated nucleic acid molecule encoding human protein 608 or a functional portion thereof. The invention further comprehends an  
20 isolated nucleic acid molecule encoding human protein 405 or a functional portion thereof. And, the invention comprehends the isolated nucleic acid molecule encoding human protein 274 or a functional portion thereof. In particular embodiments, the invention provides the isolated nucleic acid molecules identified herein by sequence numbers, as well as functional portions thereof.

- 25 The invention further encompasses a vector comprising an inventive isolated nucleic acid molecule, a composition comprising such a vector, a probe or primer which specifically hybridizes to such an isolated nucleic acid molecule, and an expression product of such an isolated nucleic acid molecule.

- The invention still further provides an isolated polypeptide herein identified as  
30 protein 608 or a functional portion thereof or a polypeptide which is at least substantially homologous or identical thereto.

The invention also provides an isolated polypeptide herein identified as protein 405 or a functional portion thereof or a polypeptide which is at least substantially homologous or identical thereto.

And, the invention provides an isolated polypeptide herein identified as  
5 protein 274 or a functional portion thereof or a polypeptide which is at least substantially homologous or identical thereto.

The invention comprehends an isolated polypeptide which is human protein 608 or a functional portion thereof, as well as an isolated polypeptide which is human protein 405 or a functional portion thereof, and an isolated polypeptide which is  
10 human protein 274 or a functional portion thereof. The invention further comprehends polypeptides identified by sequence identification numbers, as well as polypeptides from expression of nucleic acid molecules identified by sequence identification numbers; and, functional portions thereof. Further still, the invention comprehends compositions comprising an inventive polypeptide or portion thereof.  
15 Even further, the invention envisions an antibody elicited by an inventive polypeptide or a functional portion thereof, as well as a functional portion of such an antibody; and, compositions comprising such an antibody or portion thereof.

The invention further encompasses methods for preventing, treating or controlling osteoporosis or bone density or other factors causing or contributing to  
20 osteoporosis or symptoms thereof or other conditions involving mechanical stress or a lack thereof, comprising administering an inventive polypeptide or portion thereof; and accordingly, the invention comprehends uses of polypeptides in preparing a medicament or therapy for such prevention, treatment or control.

The invention even further encompasses a method for preventing, treating or  
25 controlling osteoporosis or bone density or other factors causing or contributing to osteoporosis or symptoms thereof or other conditions involving mechanical stress or a lack thereof, comprising administering an inventive vector or inventive nucleic acid molecules; and accordingly, the invention comprehends uses of such vectors or nucleic acid molecules in preparing a medicament or therapy for such prevention,  
30 treatment or control.

The invention also comprehends a method for preventing, treating or controlling osteoporosis or bone density or other factors causing or contributing to osteoporosis or symptoms thereof or other conditions involving mechanical stress or a

lack thereof, comprising administering a composition comprising a gene or functional portion thereof identified in inventive methods or an inventive model or an expression product thereof or an antibody or portion thereof elicited by such an expression product or portion thereof; and, the invention thus further comprehends uses of such genes, expression products, antibodies, portions thereof, in the preparation of a medicament or thereapy for such control, prevention or treatment.

The invention yet further provides methods for preparing a polypeptide comprising expressing the polypeptide from inventive vectors or from inventive genes or genes identified in inventive methods or models, or portions of such genes.

Further still, the invention envisions advancing research in or studies of bone development and/or osteoporosis comprising the inventive methods, materials/products, and/or models.

The invention comprehends genes differentially expressed under the influence of (a) mechanical force applied to a calvaria primary cell culture and (b) treatment of PGE2 applied to the same culture. In addition the invention involves the effect of calcium depletion. The genes that are differentially expressed are thus demonstrated to be involved in the processes that lead to osteoporosis or other mechanical stress or lack thereof related conditions.

Certain genes identified by the methods herein respond to estrogen. From the methods disclosed herein one can identify compounds to which genes identified by herein methods respond. Thus, the invention comprehends a method for affecting a gene identified by any one of the herein methods comprising contacting cells containing the gene with a compound to which the gene responds; e.g., administering the compound as a composition or formulation as herein described. Thus, for instance, with respect to genes which respond to estrogen, the invention envisions a method for affecting (e.g., stimulating expression, inhibiting expression, and the like) the gene comprising contacting a cell containing the gene with estrogen or a derivative or precursor thereof, e.g., 17-Beta estradiol and the like.

It is noted that in this disclosure, the word "comprises" can have the meaning attributed to it in U.S. Patent law; e.g., it can mean "includes".

These and other embodiments are disclosed or are obvious from and encompassed by, the following Detailed Description.

**BRIEF DESCRIPTION OF FIGURES**

The following Detailed Description, given by way of example, but not intended to limit the invention to specific embodiments described, may be understood in conjunction with the accompanying Figures, incorporated herein by reference, in which:

Figure 1A shows an absorbance profile of a fractionation of cytoplasmic RNA on a sucrose density gradient wherein the absorbance (at 254nm) is plotted against the sedimentation rate of the cytoplasmic RNA;

Figure 1B shows a purified RNA electrophoresed on an agarose gel and stained with ethidium bromide illustrating the fractionation of RNA;

Figure 2 shows a representation of DNA chip hybridization results comparing probes of total RNA (Tot) to probes derived from nuclear RNA (STP);

Figure 2A shows a table of genes identified by inventive methods, and sequences therefor or sequences of ESTs thereof (SEQ ID NOS: );

Figure 3 shows DNA and amino acid sequences for inventive nucleic acid molecule 608 and the expression product therefrom with this Figure differing from other 608 sequences herein in that Figure 3 shows additional protein sequences towards the 5' end (compare Figure 3 from about position 1025 with other 608 sequence figures herein) (SEQ ID NOS: );

Figure 4 shows the results of a 5' fragment probe of inventive nucleic acid molecule 608 on target mRNA in normal and mechanically stressed cells;

Figure 5 shows DNA and amino acid sequences for inventive nucleic acid molecule 608 and the expression product therefrom (SEQ ID NOS: );

Figure 6 shows Clustal X (1.64b) Multiple Sequence Alignment with respect to inventive nucleic acid molecule 608 and probes therefor (SEQ ID NOS: );

Figures 7 shows the results of a probe of human 405 on the target total RNA of human k562;

Figure 8 shows the results of a probe of human 405 on the target rat *cmf* RNA;

Figures 9 and 10 show the DNA and amino acid sequences for inventive nucleic acid molecule 405 and the expression product therefrom (SEQ ID NOS: );

Figure 11 shows Clustal X (1.64b) Multiple Sequence Alignment with respect to inventive nucleic acid molecule 405 and probes therefor (SEQ ID NOS: );

Figure 12 shows the results of a probe of 8 KB of human 274 on the target rat bone, rat testes and human cell line NB4 total RNA sources;

Figure 13 shows the DNA and amino acid sequences for inventive nucleic acid molecule 274 and the expression product therefrom (SEQ ID NOS: );

5 Figure 14 shows the DNA and amino acid sequences for inventive nucleic acid molecule 274 and the expression product therefrom (SEQ ID NOS: ).

(Markings on sequence figures, e.g., sequence figures such those for 608, such as dots and plus/minus signs may indicate repeats, such as IgG repeats, that may appear in many proteins; there are approximately 20 such IgG repeats in the 608 sequence listing.)

#### DETAILED DESCRIPTION

As discussed, disclosed herein is a method for identifying genes whose expression is regulated at the RNA level in an organism.

More in particular, disclosed herein is a method of identifying genes whose  
15 expression is regulated at least in part at the mRNA level by selectively stimulating an unknown target mRNA with a stress inducing element, the target mRNA being part of a larger sample. The organism may be any organism which provides suitable mRNA. The mRNA sample is derived from cellular compartments based on expression regulation and protein localization which are differentially analyzed to identify genes  
20 which are translationally regulated by the stress inducing element. This method is designed for identifying and cloning genes which are responsive to specific cues. That is, the present method is designed for identifying and cloning genes which are either up- or down- regulated responsive to a specific pathology, stress, physiological condition, and so on, and in generally to any factor that can influence cells or  
25 organisms to alter their gene expression.

This disclosure provides a novel approach to the identification and cloning of genes that are involved in fundamental cellular functions and which are regulated at any level in an organism. The basic underlying theory for this method relies on the knowledge that the regulation of gene expression can be controlled at different levels  
30 (modes) and that each different regulation level(s) is manifested by some difference in the distribution of the specific mRNAs in the cell. In genes that are regulated by translation, the mRNA is stored in the cell in an inactive form and will not be found on polysomes. Following the appropriate external cue, the mRNA is incorporated into



the polysomes and translated, and the encoded protein quickly appears. By comparing mRNA populations that are "active" or "non-active" at a given time, genes that are regulated by a mechanism referred to as the "shift mechanism" can be identified.

- 5 Genes whose main regulatory level is the active transport of mRNA from the nucleus to the cytoplasm are stored in the nucleus and at the appropriate cue the mRNA is transported to the cytoplasm. Comparison of mRNA isolated from the nucleus and cytoplasm before and after the cue can lead to the discovery of genes controlled in this way. The comparison of mRNA derived from the nucleus also
- 10 allows direct analysis of the transcription activity of many genes. For most transcriptionally activated genes a basal level of mRNA exists in the cell even when the basal transcription activity is low. Thus, increased transcription (up to five-fold) is often obscured when total cellular RNA is used for differential analysis of gene expression. The use of nuclear RNA allows direct measurement of transcription
- 15 activity of many genes, since the basal mRNA is found in the cytoplasm. The result is a major increase in sensitivity for the detection of differential expression.

In the case of mRNA stability regulation, it is expected that such mRNA would be similarly transcribed before and after cue administration, resulting in a similar abundance in nuclear mRNA pools. However, if the mRNA is stabilized

20 following the cue, its abundance in the cytoplasm would become higher. In the case of mRNA transport regulation, such mRNA is expected to exist at a high level in the nucleus and a low level in the cytoplasm prior to the cue, which situation would be reversed after administration of the cue. It is thus easy to differentiate between the two regulatory modes.

- 25 The method of the invention includes the identification of genes regulated at the translational level; genes regulated at the transcription level; genes regulated by RNA stability; genes regulated by mRNA transport rate between the nucleus and the cytoplasm; and genes regulated by differential splicing. That is, genes whose expression is at least partly controlled or regulated at the mRNA level can be
- 30 identified.

The method will identify genes encoding secreted and membrane proteins; genes encoding for nuclear proteins; genes encoding for mitochondrial proteins; and

genes encoding for cytoskeletal proteins. In addition, any other gene whose expression can be controlled at the mRNA level can be identified by this method.

As used herein, RNA refers to RNA isolated from cell cultures, cultured tissues or cells or tissues isolated from organisms which are stimulated, differentiated, exposed to a chemical compound, are infected with a pathogen or otherwise stimulated. As used herein, translation is defined as the synthesis of protein on an mRNA template.

As used herein, stimulation of translation, transcription, stability or transportation of unknown target mRNA or stimulating element, includes chemically, pathogenically, physically, or otherwise inducing or repressing an mRNA population from genes which can be derived from native tissues and/or cells under pathological and/or stress conditions. In other words, stimulating the expression of a gene's mRNA with a stress inducing element or "stressor" can include the application of an external cue, stimulus, or stimuli which stimulates or initiates translation of a mRNA stored as untranslated mRNA in the cells from the sample. The stressor may cause an increase in stability of certain mRNAs, or induce the transport of specific mRNAs from the nucleus to the cytoplasm. The stressor may also induce gene transcription. In addition to stimulating translation of mRNA from genes in native cells/tissues, stimulation can include induction and/or repression of genes under pathological and/or stress conditions. The present method utilizes a stimulus or stressor to identify unknown target genes which are regulated at the various possible levels by the stress inducing element or stressor.

The method synergistically integrates methodologies which were not previously used together.

One methodology comprises the division of cellular mRNA into separate pools of mRNA derived from polysomes, nucleus, cytoplasm or spliceosomes.

Another methodology comprises the simultaneous comparison of the relative abundance of the mRNA species found in the separate pools by a method of differential analysis such as differential display, representational difference analysis (RDA), gene expression microarray (GEM), suppressive subtraction hybridization (SSH) (Diatchenko et al., 1996), and oligonucleotide chip techniques such as the chip technology exemplified by United States Patent No. 5,545,531 to Rava et al. assigned

to Affymax Technologies N.V. and direct sequencing exemplified by WO 96/17957 patent application to Hyseq, Inc.

Briefly, subtractive hybridization is defined as subtraction of mRNA by hybridization in solution. RNAs that are common to the two pools form a duplex that can be removed, enriching for RNAs that are unique or more abundant in one pool. Differential Display is defined as reverse transcription of mRNA into cDNA and PCR amplification with degenerated primers. Comparison of the amounts amplification products (by electrophoresis) from two pools indicate transcript abundance. RDA, GEM, SSH, SAGE are described herein above.

The specific cells/tissues which are to be analyzed in order to identify translationally regulated genes, can include any suitable cells and/or tissues. Any cell type or tissue can be used, whether an established cell line or culture or whether directly isolated from an exposed organism.

The cells/tissues to be analyzed under the present method are selectively stimulated or "stressed" utilizing a physiological, chemical, environmental and/or pathological stress inducing element or stressor, in order to stimulate the translation of mRNA within the sample tissue and identify genes whose expression is regulated at least in part at the mRNA level. Stimulation can cause up or down regulation. Following stimulation, RNA is isolated or extracted from the cells/tissues. The isolation of the RNA can be performed utilizing techniques which are well known to those skilled in the art and are described, for example, in "Molecular Cloning; A Laboratory Manual" (Cold Springs Harbor Laboratory Press, Cold Spring Harbor, New York, 1989). Other methods for the isolation and extraction of RNA from cells/tissue can be used and will be known to those of ordinary skill in the art. (Mach et al., 1986, Jefferies et al., 1994). However, may variations of these methodologies have been published. The methods described herein were carefully selected after many trials.

The mRNAs which are actively engaged in translation and those which remain untranslated can be separated utilizing a procedure such as fractionation on a sucrose density gradient, high performance gel filtration chromatography, or polyacrylamide gel matrix separation (Ogishima et al., 1984, Menaker et al., 1974, Hirama et al., 1986, Mechler, 1987, and Bharucha and Murthy, 1992), since mRNAs that are being translated are loaded with ribosomes and, therefore, will migrate differently on a

density gradient than ribosome-free untranslated mRNAs. By comparing mRNA populations that are active or non-active in translation at a given time, genes that are regulated by the "shift mechanism" can be identified.

Polysomal fractionation and specific analysis can be facilitated by treatment of  
5 target cell/tissue with drugs that will specifically inhibit or modulate transcription or translation. Examples of such drugs are actinomycin D and cyclohexamide, respectively.

The fractionation can be completed to create polysomal subdivisions. The subdivisions can be made to discriminate between total polyribosomes or membrane  
10 bound ribosomes by methods known in the art (Mechler, 1987). Further, the mRNA sample can additionally be fractionated into one or more of at least the following subsegments or fractions: cytoplasmatic, nuclear, polyribosomal, sub polyribosomal, microsomal or rough endoplasmic reticulum, mitochondrial and spliceosome associated mRNA by methods known in the art.

15 More specifically, nuclear fractions can be obtained using the method set forth in the article entitled Abundant Nuclear Ribonucleoprotein Form of CAD RNA (Sperling, 1984) as set forth in the Examples, thus allowing nuclear RNA to be utilized for a method of identifying genes which are regulated or responsive to stress conditions. Further, antisense RNA can be utilized as a method for identifying genes  
20 which are responsive to specific pathology or stress conditions. Antisense RNA can be isolated using the methods described by Dimitrijevic, whose abstract details the methods utilized for obtaining and isolating antisense RNA from a sample. Additionally, microsomal fractions may be obtained using the methods of the present invention as set forth in the Experimental Section which are modifications of the  
25 methods disclosed by Walter and Blobel in 1983.

Following isolation and division of the total mRNA population into separate expression regulation and protein localization pools of mRNA, the relative abundance of the many mRNA species found in these pools are simultaneously compared using a differential analysis technique such as differential display, oligonucleotide chips,  
30 representational difference analysis (RDA), GEM-Gene Expression Microarrays (Scheda et al., 1995, Aiello et al., 1994, Shen et al., 1995, Bauer et al., 1993, Liang and Pardee, 1992, Liang and Pardee, 1995, Liang et al., 1993, Braun et al., 1995, Hubank and Schatz, 1994) and suppressive subtraction hybridization (SSH). The

RNA isolated from the fractions can be further purified into mRNA without the ribosomal RNA by poly A selection. It should be noted that multiple pools can be analyzed utilizing this method. That is, different cell aliquots subjected to different stressors can be compared with each other as well as with the reference sample.

- 5        Labeled nucleic acid probes (in a cDNA, PCR product or rRNA transcribed from the cDNA) made from RNA derived from polysomal, non-polysomal, mRNPs, nuclear, cytoplasmic, or spliceosome fractions can be used as probes, to identify clones of cDNA, genomic clones, and mRNA species that are fixed onto a solid matrix-like microarrays such as (GEM), that shown in United States Patent Number
- 10    5,545,531 to Rava et al. and WO96/17957 to Hyseq, Inc., and membranes of any kind where clones can be either blotted after electrophoresis or directly loaded (dot blot) onto the membrane. The label can be radioactive, fluorescent, or incorporating a modified base such as digoxigenin and biotin.

- Comparison between the fractions derived from the polysomal or
- 15    polyribosomal fraction or other fractions to the total unfractionated material is essential to discriminate between differentials in expression levels that are the result of transcription modulation from those that result from modulation of translation per se. The polysomal fractions or groups can include membrane bound polysomes, loose or tight polysomes, or free unbound polysome groups.

- 20    The importance of utilizing the polysomal sub-population in order to identify differentially (translationally) expressed genes is shown in Example 1 where a number of genes were not detected as translationally expressed under heat shock inducement when total mRNA was used as the detection probe but, however, when polysomal mRNA was used as a probe, a number of genes were identified as
- 25    differentially expressed. As shown in Example 1, a number of genes under heat shock inducement with total mRNA derived probe were detected when probed with polysomal mRNA fractions. Heat shock, being a model for acute diseases such as ischemic diseases, reveal the importance of the polysomal probe. Cells store critical mRNAs in an inactive form so that in an acute situation they can be quickly loaded
- 30    onto polysomes (without the need to wait for their production by transcription) and translated to produce the proteins the cells require for their survival under stress.

The present method for identifying translationally regulated genes is not limited by the source of the mRNA pools. Therefore, the present method can be

utilized to clone genes from native cells/tissue under pathological and/or stress conditions that are regulated by the "shift mechanism," as well as genes that are induced/repressed under pathological and/or stress conditions. Pathologies can include disease states including those diseases caused by pathogens and trauma.

- 5 Stress conditions can also include disease states, physical and psychological trauma, and environmental stresses. Following analysis by the selected method of differential analysis, the genes which have been identified as being regulated by translation can be cloned by any suitable cloning methodologies known to those skilled in the art. (Lisitsyn and Wigler, 1993).

- 10 Differential comparisons can be made of all possible permutations of polysomal vs. non-polysomal RNA where the definition of the fraction type is done, for example, by absorbance profile at 254nm, density of the sucrose gradient as shown in Figure 1A (or another size standard if high pressure liquid chromatography or gel systems are used) and types of RNA that are stained with ethidium bromide  
15 after electrophoresis of the fractions on agarous gels are completed, as shown in Figure 1B. In Figure 1A, the polysomal fractions are those that have mRNA with more than two ribosomes loaded. The materials and methods for this comparison are set forth below in the experimental section.

- Differential comparisons can also include polysomal vs. non-polysomal  
20 fractions in each condition. By "condition" it is meant that cells from the same source, such as a cell line, a primary cell, or a tissue that undergoes different treatment or has been modified to have different features or to express different sets of genes. For example, this can be accomplished by differentiation, transformation, application of the stress such as oxygen deprivation, chemical treatment, or radiation.

- 25 Permutations can include, for example:

1. polysomal fractions between conditions individually (migrating in the same density) or in a pool;
2. non-polysomal fractions between conditions individually (migrating in the same density) or in a pool;
- 30 3. non-polysomal to polysomal between conditions and within each condition individually (migrating in the same density) or in a pool; and

4. each of the fractions being polysomal and non-polysomal individually (migrating in the same density) or in a pool that can be compared to total RNA that is unfractionated.

The method described above for the identification of genes regulated on the translational level has a number of applications. A particular application for this method is its use for the detection of changes in the pattern of mRNA expression in cells/tissue associated with any physiological or pathological change. By comparing the translated versus untranslated mRNAs, the effect of the physiological or pathological cue or stress on the change of the pattern of mRNA expression in the cell/tissue can be observed and/or detected. This method can be used to study the effects of a number of cues, stimuli, or stressors to ascertain their effect or contribution to various physiological and pathological activities of the cell/tissue. In particular, the present method can be used to analyze the results of the administrations of pharmaceuticals (drugs) or other chemicals to an individual by comparing the mRNA pattern of a tissue before and after the administration of the drug or chemical. This analysis allows for the identification of drugs, chemicals, or other stimuli which affect cells/tissue at the level of translational regulation. Utilizing this method, it is possible to ascertain if particular mRNA species are involved in particular physiological or disease states and, in particular, to ascertain the specific cells/tissue wherein the external stimulus, i.e., a drug, affects a gene which is regulated at the translational level.

The identification of a subgroup of genes regulated on the translational level involved a method for identifying gene sequences coding for internal ribosome entry sites (IRES), including the general steps of inhibiting 5'cap-dependant mRNA translation in a cell, collecting a pool of mRNA from the cells, and differentially analyzing the pool of mRNA to identify genes with sequences coding for internal ribosome entry sites. The inhibiting step can be further defined as selecting for non-5'-cap dependent mRNA translation or by incorporating a gene, such as a gene coding for a protease such as poliovirus 2A protease. The method can include the step of controlling the expression of the gene. The analyzing step can be further defined as differential display analysis, or as representational difference analysis, or as performing a gene expression microarray analysis. The method can include the further step of cloning genes identified as being translationally regulated. The

analyzing step can distinguish between polysomal fractions that migrate in the same density individually or in a pool. The analyzing step can distinguish between nonpolysomal fractions individually or as a pool. The analyzing step can distinguish between stimulated polysomal and nonpolysomal fractions individually or in a pool.

- 5 And, the analyzing step can distinguish between each of the polysomal and nonpolysomal fractions individually or in a pool compared to an unfractionated total RNA pool.

Utilizing these methods, it is possible to ascertain if particular mRNA species are involved in particular physiological or disease states and, in particular, to ascertain  
10 the specific cells/tissue wherein the external stimulus, e.g., a drug, affects a gene which is regulated at the translational level.

- Accordingly, in an aspect, the application also discloses a method for determining risk of developing a physiological or disease state based upon absence or decrease from normal cells of mRNA or protein from a gene shown to be down  
15 regulated by the inventive or herein disclosed methods in a mammal comprising:
- (a) determining the level or status of mRNA in cells, e.g., bone cells (for instance, osteoblasts and/or osteoclasts) of said mammal; and/or
  - (b) determining the level or status of corresponding protein in cells of said mammal; and
  - 20 (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity
  - 25 of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining an absence or decrease from normal cells of mRNA or protein and thus risk.

- In another aspect, the disclosure herein provides a method for determining risk of physiological or disease state based upon presence or increase from normal cells of  
30 mRNA or protein from a gene shown to be upregulated by the inventive or herein disclosed methods in a mammal comprising:
- (a) determining the level or status of mRNA in cells, e.g., bone cells (for instance, osteoblasts and/or osteoclasts) of said mammal; and/or



(b) determining the level or status of corresponding protein in cells of said mammal; and

- (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining a presence or increase from normal cells of mRNA or protein and thus risk.

The foregoing methods can be employed in inventive methods for testing a medicament for or a gene therapy approach to a physiological or disease state or other factors causing or contributing thereto or to symptoms thereof based on absence or decrease from normal cells or presence or increase from normal cells of mRNA or protein of identified genes additionally comprising: (a') administering the medicament or the gene therapy; and from the comparing determining an absence or decrease from normal cells or presence or increase from normal cells of the relevant mRNA or protein and thus efficacy of the medicament or the gene therapy.

Similarly, in still further aspects, the disclosure herein provides methods for treating, preventing or controlling a physiological or disease state comprising administering a medicament or treatment therefor or for a cause thereof or a symptom thereof, including the foregoing detection methods. For instance, from the comparing one determines an absence or decrease from normal cells or presence or increase from normal cells of particular mRNA or protein and thus risk and administers a the medicament or treatment.

The methods can additionally comprise using the steps in conjunction with another test method akin to those described above, e.g., having a same or similar preamble recitation and comprising:

- (d) determining the level or status of a second gene mRNA in cells of said mammal; and/or
- (e) determining the level or status of protein expressed by a second gene product in cells of said mammal; and

(f) comparing said level or status of that mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein.

The absence or decrease or presence or increase may be correlated to risk.

Thus, the second gene can be identified by methods of the invention or as disclosed herein. Or alternatively or additionally, the second gene and/or the additional steps can be determined in accordance with other methods, e.g., other methods for determining the risk of the physiological or disease state or a condition or factor associated therewith. Thus, such methods can be used in conjunction with methods herein to advance or improve diagnostic or detection methodologies.

In the methods, steps (a) and/or (b) and optionally (d) and/or (e) are carried out *in vivo* and/or steps (a) and/or (b) and optionally (d) and/or (e) are carried out *in vitro*.

The determination in step (a) and optionally in step (d) can be effected by employing

- (i) a nucleic acid sequence corresponding to at least a part of the gene encoding at least part of the protein and optionally a second nucleic acid sequence corresponding to at least a part of the second gene encoding at least part of the second protein;
- (ii) a nucleic acid sequence complementary to the nucleic acid sequence(s) of (i); or
- (iii) a primer or a primer pair hybridizing to the nucleic acid sequence(s) of (i) or (ii).

The determination in step (b) and optionally of step (e) can be effected by employing an antibody or a fragment thereof that specifically binds to the protein and optionally by employing a second antibody or a fragment thereof which specifically binds to the second protein.

Cells may be considered "normal" in the methods by having an absence of the physiological or disease being tested for; or by any other standard definition recognized in the art.

The medicament or treatment can be any conventional medicament or treatment for the physiological or disease. Alternatively or additionally, the medicament or treatment can be the particular protein of the gene detected in the inventive methods or a functional portion thereof, or that which inhibits that protein, e.g., binds to it. Similarly, additionally or alternatively, the medicament or treatment can be a vector which expresses the protein of the gene detected in the inventive methods or a functional portion thereof or that which inhibits expression of that gene; again, for instance, that which can bind to it and/or otherwise prevents its transcription or translation. The selection of administering a protein or that which expresses it, or of administering that which inhibits the protein or the gene expression, can be done without undue experimentation, e.g., based on down regulation or up regulation as determined by inventive methods (e.g., in the osteoporosis model).

In an aspect of the invention, the stimulus in inventive methods is mechanical stress or a lack thereof, e.g., with respect to bone cells which retain their characteristic thereof in cultures.

In a further aspect, the invention provides an application of inventive methods with respect to osteoporosis, a major health problem; and provides inventive products and uses therefor. As discussed, osteoporosis or porous bone, is a progressive and chronic disease characterized by low bone mass and structural deterioration of bone tissue, with bone loss being possibly without symptoms, leading to bone fragility and an increased susceptibility to fractures of the hip, spine, and wrist (diminishing bone strength).

Osteoporosis is histologically, biochemical and kinetically heterogeneous. Data points to causes such as: deficiency of estrogen and deficiency of calcium.

Calcium is an essential nutrient that is involved in most metabolic processes and the phosphate salts of which provide mechanical rigidity to the bones and teeth, where 99% of the body's calcium resides. The calcium in the skeleton has the additional role of acting as a reserve supply of calcium to meet the body's metabolic needs in states of calcium deficiency. Calcium deficiency is easily induced because of the obligatory losses of calcium via the bowel, kidneys, and skin. Calcium deficiency delays the consolidation of the skeleton, may cause mobilization of bones and has been shown in animals to lead to osteoporosis.

Further, bone is composed of a collagen-rich organic matrix impregnated with mineral, largely calcium and phosphate. Two major forms of bone exist, compact cortical bone forms the external envelopes of the skeleton and trabecular or medullary bone forms plates that traverse the internal cavities of the skeleton. The responses of these two forms to metabolic influences and their susceptibility to fracture differ. Bone undergoes continuous remodeling (turnover) throughout life. Osteoclasts are the cells in the skeleton that responsible for breaking down bones, osteoblasts on the other hand, are capable of forming new bones. Mechanical and electrical forces, hormones and local regulatory factors influence remodeling. Peak bone mass is mainly genetically determined, though dietary factors and physical activity can have positive effects. Peak bone mass is attained at the point when skeletal growth ceases, after which time bone loss starts. Bone mass declines throughout life due to an imbalance in this process.

It is noted that the World Health Organization (WHO Technical Report Series: 843, 1994) characterizes "normal", e.g., as to women, as bone mineral density (BMD) or bone mineral content (BMC) that is greater than or equal to 1 standard deviation (SD) below the young adult reference range; "low bone mass" as BMD or BMC 1-2.5 SD below the mean of young healthy adults, e.g., women; "osteoporosis" as BMD or BMC greater than 2.5 SD below the mean of young healthy adults, e.g., women; and "severe osteoporosis" as BMD or BMC greater than 2.5 SD below the mean of young healthy adults, e.g., women and the presence of one or more fragility fractures. From this information and the knowledge in the art, the skilled artisan can determine and employ "normal" cells, without any undue experimentation.

Osteoblasts are particularly sensitive to aging phenomena--more sensitive than are osteoclasts--so the negative bone balance increases with increasing age. Age-dependent bone loss is aggravated by reduced calcium absorption, a mutation in the collagen gene and polymorphism in TGF-beta and estrogen receptor proteins.

Cells bind to ECM (extracellular matrix) via specific cell surface receptors such as integrins. When engaging with ECM ligands, these receptors can activate signal transduction pathways within the cells and may act as mechanochemical transducers. Thus, interaction of cells with ECM can modulate gene expression. Among the genes that are, in part, controlled by cell-ECM interactions are those for

certain ECM components themselves. Bone cells, remodel their matrix and reorient bone trabeculae in response to mechanical strain.

- Accordingly, the nature of the bone cell response can relate to the state of differentiation. Furthermore, evidence shows that prostaglandins are likely to play an important role in the physiologic and pathologic responses of bone tissue. Prostaglandins can stimulate and inhibit bone resorption and formation. Prostaglandins mediate bone loss due to immobilization, but prostaglandin E2 (PGE2) stimulates bone formation *in vivo*. Prostaglandin production by bone cells is highly regulated by mechanical forces, cytokines, growth factors and systemic hormones.
- 10 Mechanical stimulation applied to cultured bone cells results in increased production of several prostaglandins including PGE2, prostaglandin I2 (PGI2), and prostaglandin F2a. Addition of indomethacin, which blocks endogenous prostaglandin production, neutralizes the effect of mechanical stress treatment.

- Cells isolated from calvaria bone maintain their osteoblastic phenotype in culture. Genetics factors play an important role in the pathogenesis of osteoporosis. It is suggested that up to 85%-90% of the variance in bone mineral density is genetically determined. Thus, calvaria bone cells were used in methods of the invention. Genes differentially expressed under the influence of (a) mechanical force applied to a calvaria primary cell culture and (b) treatment of PGE2 applied to the same culture. In
- 15 addition, the effect of calcium depletion is also shown. The genes that result differentially expressed are thus demonstrated to be involved in the processes that lead to osteoporosis, and ergo osteoporosis.

- It is well accepted that the main process that is characteristic of osteoporosis – enhanced bone resorption – takes place not only in conditions of low estrogen production (menopause women), but also in some other conditions, like treatment with glucocorticoids or bone immobilization. Therefore, it was reasoned that application of mechanical force is stimulatory for bone formation. To model this process, as discussed in the Examples, e.g., Example 2, primary rat calvaria cells grown on elastic membranes and stretched together with this membrane for 20
- 20 minutes. Genes expression patterns were compared before and after the application of mechanical force. Particular genes were found to be differentially regulated and/or differentially expressed following mechanical stimulation, validating the osteoporosis model; and, showing that the inventive methods can be used to identify genes,

expression products therefrom, probes/primers for such genes, as well as uses for such genes, expression products, probes/primers, *inter alia*.

In an aspect the invention provides a gene identification process. Steps involved in the gene identification process comprise one or more or all of the following: Preparation of probes from the model system (mechanical force, calvaria primary culture); analysis of DNA chip hybridization; sequencing of clones showing differential expression; and full-length cloning of clones of interest (cloning can be by a variety of known methodologies).

In yet another aspect the invention provides an osteoporosis model or a model for other conditions caused by mechanical stress or force, e.g., bone mass formation, comprising rat calvaria cells or another cell which retains osteoblast or osteoclast nature in cell cultures, being subjected to mechanical or other bone growth/formation inducing stress or stimuli or bone loss inducing stress or stimuli.

With respect to mechanical stress and osteoporosis, it is well documented that exercise has very beneficial effects on bone mass. The effect of the zero of gravity on astronauts and their need to do a lot of exercise is also believed well known.

However, as far as the inventors know, efforts to isolate genes involved in the biological interpretation of mechanical stress signals into increase in bone mass have not heretofore been done. And, mention is made of: Binderman I, Duksin D, Harell A,

Katzir E, Sachs L. Formation of bone tissue in culture from isolated bone cells. J Cell Biol 1974 May;61(2):427-39; Harell A, Dekel S, Binderman I. Biochemical effect of mechanical stress on cultured bone cells. Calcif Tissue Res 1977 May;22 Suppl:202-7; Somjen D, Binderman I, Berger E, Harell A. Bone remodelling induced by physical stress is prostaglandin E2 mediated. Biochim Biophys Acta 1980 Jan 3;627(1):91-100; Shimshoni Z, Binderman I, Fine N, Somjen D. Mechanical and hormonal

stimulation of cell cultures derived from young rat mandible condyle. Arch Oral Biol 1984;29(10):827-31; Binderman I, Shimshoni Z, Somjen D. Biochemical pathways involved in the translation of physical stimulus into biological message. Calcif Tissue Int 1984;36 Suppl 1:S82-5; Binderman I, Zor U, Kaye AM, Shimshoni Z, Harell A, Somjen D. The transduction of mechanical force into biochemical events in bone cells may involve activation of phospholipase A2. Calcif Tissue Int 1988 Apr;42(4):261-6; Binderman I, Berger E, Fine N, Shimshoni Z, Harell A, Somjen D. Calvaria derived

osteogenic cells: phenotypic expression in culture. *Connect Tissue Res* 1989;20(1-4):41-7.

The inventive osteoporosis model or model for other conditions caused by mechanical stress or force or lack thereof, encompasses inventive methods and products of inventive methods employed under conditions of little or no gravity, e.g., the results of performing inventive methods, such as those exemplified or analogous to those exemplified, for instance without applying mechanical stress and under conditions of little or no gravity such as on a space vehicle such as a Space Shuttle or a space station as is being constructed.

The invention is still further aspects provides CMF274, expression products therefrom, probes/primers therefor, and uses for such gene, expression products and primers/probes, as well as of functional portions of the gene or of the expression product.

The invention is still further aspects provides CMF405, expression products therefrom, probes/primers therefor, and uses for such gene, expression products and primers/probes, as well as of functional portions of the gene or of the expression product.

The invention is still further aspects provides CMF608, expression products therefrom, probes/primers therefor, and uses for such gene, expression products and primers/probes, as well as of functional portions of the gene or of the expression product.

Species of origin of the sequences: all initial sequences (short fragments of ~500bp) were rat. For 405 a homologue in the form of a partially characterized mRNA was found but there is no published information on its expression in bones. The study in rats was done to prove its function and possible uses in humans (directly or indirectly) for therapeutics. Once the rat sequence is known the isolation of the human homologues can be within the ambit of the skilled artisan; and thus, this disclosure is intended to cover the human homologues as well because these homologues fall within a degree of homology included within the present invention.

More in particular, respect to the herein mentioned nucleic acid molecules and polypeptides therefrom, e.g., the aforementioned nucleic acid molecules (608, 405, 274) and polypeptides expressed from them, the invention further comprehends isolated and/or purified nucleic acid molecules and isolated and/or purified

polypeptides having at least about 70%, preferably at least about 75% or about 77% identity or homology ("substantially homologous or identical"), advantageously at least about 80% or about 83%, such as at least about 85% or about 87% homology or identity ("significantly homologous or identical"), for instance at least about 90% or about 93% identity or homology ("highly homologous or identical"), more  
5 advantageously at least about 95%, e.g., at least about 97%, about 98%, about 99% or even about 100% identity or homology ("very highly homologous or identical" to "identical"; or from about 84-100% identity considered "highly conserved"). The invention also comprehends that these nucleic acid molecules and polypeptides can be  
10 used in the same fashion as the herein or aforementioned nucleic acid molecules and polypeptides.

Nucleotide sequence homology can be determined using the "Align" program of Myers and Miller, ("Optimal Alignments in Linear Space", CABIOS 4, 11-17, 1988, incorporated herein by reference) and available at NCBI. Alternatively or  
15 additionally, the term "homology" or "identity", for instance, with respect to a nucleotide or amino acid sequence, can indicate a quantitative measure of homology between two sequences. The percent sequence homology can be calculated as  $(N_{ref} - N_{dif}) * 100 / N_{ref}$ , wherein  $N_{dif}$  is the total number of non-identical residues in the two sequences when aligned and wherein  $N_{ref}$  is the number of residues in one of the  
20 sequences. Hence, the DNA sequence AGTCAGTC will have a sequence similarity of 75% with the sequence AATCAATC ( $N_{ref} = 8$ ;  $N_{dif} = 2$ ).

Alternatively or additionally, "homology" or "identity" with respect to sequences can refer to the number of positions with identical nucleotides or amino acids divided by the number of nucleotides or amino acids in the shorter of the two  
25 sequences wherein alignment of the two sequences can be determined in accordance with the Wilbur and Lipman algorithm (Wilbur and Lipman, 1983 PNAS USA 80:726, incorporated herein by reference), for instance, using a window size of 20 nucleotides, a word length of 4 nucleotides, and a gap penalty of 4, and computer-assisted analysis and interpretation of the sequence data including alignment can be  
30 conveniently performed using commercially available programs (e.g., Intelligenetics™ Suite, Intelligenetics Inc. CA). When RNA sequences are said to be similar, or have a degree of sequence identity or homology with DNA sequences, thymidine (T)



in the DNA sequence is considered equal to uracil (U) in the RNA sequence (*see also* alignment used in Figures).

RNA sequences within the scope of the invention can be derived from DNA sequences, by thymidine (T) in the DNA sequence being considered equal to uracil (U) in RNA sequences.

Additionally or alternatively, amino acid sequence similarity or identity or homology can be determined using the BlastP program (Altschul *et al.*, Nucl. Acids Res. 25, 3389-3402, incorporated herein by reference) and available at NCBI. The following references (each incorporated herein by reference) provide algorithms for comparing the relative identity or homology of amino acid residues of two proteins, and additionally or alternatively with respect to the foregoing, the teachings in these references can be used for determining percent homology or identity: Needleman SB and Wunsch CD, "A general method applicable to the search for similarities in the amino acid sequences of two proteins," J. Mol. Biol. 48:444-453 (1970); Smith TF and Waterman MS, "Comparison of Bio-sequences," Advances in Applied Mathematics 2:482-489 (1981); Smith TF, Waterman MS and Sadler JR, "Statistical characterization of nucleic acid sequence functional domains," Nucleic Acids Res. 11:2205-2220 (1983); Feng DF and Doolittle RF, "Progressive sequence alignment as a prerequisite to correct phylogenetic trees," J. of Molec. Evol. 25:351-360 (1987); Higgins DG and Sharp PM, "Fast and sensitive multiple sequence alignment on a microcomputer," CABIOS, 5: 151-153 (1989); Thompson JD, Higgins DG and Gibson TJ, "ClusterW: improving the sensitivity of progressive multiple sequence alignment through sequence weighing, positions-specific gap penalties and weight matrix choice," Nucleic Acid Res. 22:4673-480 (1994); and, Devereux J, Haeblerle P and Smithies O, "A comprehensive set of sequence analysis program for the VAX," Nucl. Acids Res. 12: 387-395 (1984).

In this fashion, by comprehending nucleic acid molecules and polypeptides having such homology to the particular sequences disclosed, it is envisioned that the invention encompasses human and other homologues to the disclosed sequences, within the herein terms. Identification and/or isolation of corresponding human sequences can be any suitable method, for instance, by analysis of hybridization of herein defined genes (such as genes identified herein and/or identified by inventive methods herein) or suitable portions thereof, e.g., primers/probes derived from herein

defined genes; for instance, in PCR amplification of portions of the human genome by such primers/probes and/or labeled hybridization analysis of herein defined genes or portions thereof to portions of the human genome (*see also* discussions *infra*, e.g., concerning PCR, hybridization, *inter alia*).

- 5 Furthermore, by comprehending proteins having homology to the gene products of genes identified herein (e.g., 405, 608, 274) as well as of genes identified by the methods disclosed herein, the invention comprehends proteins which are "functional" proteins derived from gene products identified herein, as well as from gene products of genes identified by the methods disclosed herein; e.g., truncated
- 10 forms of proteins identified herein or expressed by genes identified by the methods disclosed herein.

- As to uses, the inventive genes and expression products as well as genes identified by the herein disclosed methods and expression products thereof (including "functional" variations of such expression products, and ergo truncated portions of
- 15 herein defined genes such as portions of herein defined genes which encode a functional portion of an expression product) are useful in treating, preventing or controlling or diagnosing or observing or studying osteoporosis or processes thereof or mechanical stress conditions or absence or reduced mechanical stress conditions. They may aid in bone density. They may be useful for diagnostic purposes. They
- 20 may be used for determining predisposition to high or low bone density or for determining gene association or other factors associated with high bone mass or low bone mass.

- For instance, 608 expression causes cells to differentiate into osteoblasts and chondrocytes. The expression product of 608, or if cells or vectors expressing 608
- 25 may cause cells to selectively differentiate and thereby increase or alter bone density. Detecting levels of 608 mRNA or expression and comparing it to "normal" non-osteopathic levels may allow one to detect who may be at risk for osteoporosis or lower levels of osteoblasts and chondrocytes.

- 405 expression impacts upon bone density by being characteristic for
- 30 osteogenic and chondrogenic cells in their differentiation preceeding matrix calcification. The expression product of 405 or cells or vectors expressing it may cause cells to differentiate into osteogenic and chondrogenic cells and thereby increase matrix calcification and bone density. One may detect a risk of low bone

density or low matrix calcification or osteoporosis by determining levels of expression of 405 or of mRNA and compare it to "normal" levels.

274 is implicated in lymphoid precursors in bone marrow. Under expression may lead to less lymphoid cells and bones that are more susceptible to environmental factors or other than genetic factors of osteoporosis, e.g., cancer causes of osteoporosis. One may detect a risk of predisposition of bones towards susceptibility to environmental factors, or less lymphoid cells, or osteoporosis by determining levels of expression of 405 or of mRNA and comparing it to "normal" levels.

Further, genes which were upregulated and identified by the method of the present invention are of interest. That which may inhibit these genes and/or the expression products therefrom or portions thereof, e.g., antibodies or functional portions thereof or other compounds which bind thereto, may be useful in preventing, controlling or treating osteoporosis or factors leading thereto or causing osteoporosis or other conditions involving mechanical stress or a lack thereof, and the genes may be targets for anti-osteoporosis treatment or therapy, as well as for study of osteoporosis or factors leading thereto or causes thereof, e.g., determining predisposition to high or low bone density or for determining gene association or other factors associated with high bone mass or low bone mass.

Among these, three identified upregulated RGD-containing proteins, ADAMTS-1 and complement 3 itself (potential prevention of osteoclast attraction) and two proteins of the SARP family (secreted apoptosis related proteins) as potential modifiers of programmed cell death in bone formation were identified.

Similarly, genes which were downregulated and identified by the method of the present invention are interesting. These genes and/or the expression products therefrom and/or a functional portion thereof may be useful in preventing, controlling or treating osteoporosis or factors leading thereto or causing osteoporosis, or other conditions involving mechanical stress or a lack thereof, and the genes may be targets for anti-osteoporosis treatment or therapy, as well as for study of osteoporosis or factors leading thereto or causes thereof, e.g., determining predisposition to high or low bone density or for determining gene association or other factors associated with high bone mass or low bone mass.

Accordingly, in an aspect, the invention provides a method for determining risk of developing osteoporosis or low or high bone density or other factors causing or

contributing to osteoporosis or other conditions involving mechanical stress or a lack thereof, based upon absence or decrease from normal cells of mRNA or protein from a gene shown to be down regulated by the inventive methods in a mammal comprising:

- 5 (a) determining the level or status of mRNA in cells, e.g., bone cells (for instance, osteoblasts and/or osteoclasts) of said mammal; and/or
- (b) determining the level or status of corresponding protein in cells of said mammal; and
- (c) comparing said level or status of mRNA and/or protein with the corresponding
- 10 level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including
- 15 post-translational modifications of the protein, and from the comparing determining an absence or decrease from normal cells of mRNA or protein and thus risk.

In another aspect, the invention provides a method for determining risk of developing osteoporosis or low or high bone density or other factors causing or contributing to osteoporosis or other conditions involving mechanical stress or a lack thereof, based upon presence or increase from normal cells of mRNA or protein from a gene shown to be upregulated by the inventive methods in a mammal comprising:

- (a) determining the level or status of mRNA in cells, e.g., bone cells (for instance, osteoblasts and/or osteoclasts) of said mammal; and/or
- (b) determining the level or status of corresponding protein in cells of said
- 25 mammal; and
- (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a
- 30 mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining a presence or increase from normal cells of mRNA or protein and thus risk.

As mentioned, 608 expression causes cells to differentiate into osteoblasts and chondrocytes. Thus, in a further aspect, the invention provides a method for determining risk of developing osteoporosis or low or high bone density or other factors causing or contributing to osteoporosis or lower levels of osteoblasts and chondrocytes or other conditions involving mechanical stress or a lack thereof, based upon absence or decrease from normal cells of mRNA or protein from 608 comprising:

- (a) determining the level or status of mRNA in cells, e.g., bone cells (for instance, osteoblasts and/or osteoclasts) of said mammal; and/or
- 10 (b) determining the level or status of corresponding protein in cells of said mammal; and
- (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a
- 15 transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining a presence or increase from normal cells of mRNA or protein and thus risk.

- 20 As discussed herein, 405 expression impacts upon bone density by being characteristic for osteogenic and chondrogenic cells in their differentiation preceeding matrix calcification. Accordingly, in a still further aspect, the invention provides a method for determining risk of developing osteoporosis or low or high bone density or other factors causing or contributing to osteoporosis or of imbalance as to
- 25 osteogenic and chondrogenic cells or other conditions involving mechanical stress or a lack thereof, based upon absence or decrease from normal cells or presence or increase from normal cells, e.g., absence or decrease from normal cells of mRNA or protein from 405 in a mammal comprising:

- (a) determining the level or status of mRNA in cells, e.g., bone cells (for instance, osteoblasts and/or osteoclasts) of said mammal; and/or
- 30 (b) determining the level or status of corresponding protein in cells of said mammal; and

(c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining an absence or decrease from normal cells of mRNA or protein and thus risk.

Likewise, as discussed herein, 274 is implicated in lymphoid precursors in bone marrow.

Therefore, in yet another aspect, the invention provides a method for determining risk of developing osteoporosis or low or high bone density or other factors causing or contributing to osteoporosis or of being susceptible to environmental factors or other than genetic factors of osteoporosis, e.g., cancer causes of osteoporosis or of predisposition of bones towards susceptibility to environmental factors, or less lymphoid cells, or osteoporosis, or other conditions involving mechanical stress or a lack thereof, based upon presence or increase from normal cells or absence or decrease from normal cells, e.g., absence or decrease from normal cells of mRNA or protein from 274 in a mammal comprising:

- (a) determining the level or status of mRNA in cells, e.g., bone cells (for instance, osteoblasts and/or osteoclasts) of said mammal; and/or
- (b) determining the level or status of corresponding protein in cells of said mammal; and
- (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining an absence or decrease from normal cells of mRNA or protein and thus risk.

The foregoing methods, in still further aspects of the invention, can be employed in inventive methods for testing a medicament for or a gene therapy

approach to osteoporosis or bone density or other factors causing or contributing to osteoporosis or symptoms thereof or other conditions involving mechanical stress or a lack thereof, based on absence or decrease from normal cells or presence or increase from normal cells of mRNA or protein of identified genes additionally comprising:

- 5 (a') administering the medicament or the gene therapy; and from the comparing determining an absence or decrease from normal cells or presence or increase from normal cells of the relevant mRNA or protein and thus efficacy of the medicament or the gene therapy.

- Similarly, in still further aspects, the invention provides methods for treating,  
10 preventing or controlling osteoporosis or other conditions involving mechanical stress or a lack thereof, comprising administering a medicament or treatment for osteoporosis or a cause thereof or a symptom thereof, including the foregoing detection methods. For instance, from the comparing one determines an absence or decrease from normal cells or presence or increase from normal cells of particular  
15 mRNA or protein and thus risk and administers a the medicament or treatment.

- The cells in the inventive methods can be *in vitro* or *in vivo* or from any suitable mammal, e.g., a human, a domesticated animal, for instance a companion animal or livestock, or a laboratory animal, such as a rat, mouse or the like; and, the cells can be from any stage of the mammal's development, such as embryonic, mature  
20 or adult, immature or child, newborn, or elderly, and the like.

- It is noted that as to CMF608, the inventors did not see any differences in its expression between normal and ovariectomized rats suggesting it may not necessarily *per se* be a marker for bone intensity. Similarly, for CMF405 there was no change in expression after ovariectomy. In addition, expression in a few non-bone tissues  
25 suggests it may not necessarily *per se* be a marker.

Thus, inventive methods can additionally comprise using the steps in conjunction with another test method akin to those described herein, e.g., having a same or similar preamble recitation and comprising:

- (d) determining the level or status of a second gene mRNA in bone cells of said  
30 mammal; and/or  
(e) determining the level or status of protein expressed by a second gene product in bone cells of said mammal; and

- (f) comparing said level or status of that mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein.

The absence or decrease or presence or increase detected may be correlated to risk. Thus, the second gene can be identified by methods of the invention. Or, alternatively or additionally, the second gene and/or the additional steps can be determined in accordance with other methods, e.g., as in U.S. Patents Nos. 5,834,200 and/or 5,691,153.

Likewise, it is within the invention that the inventive genes or genes identified by inventive methods herein or portions thereof can be the subject of other or analogous methods, such as a method for determining predisposition to high or low bone density comprising detecting the under or over expression of the gene or abnormalities in a receptor for a gene product or polymorphysim; *see, e.g.*, U.S. Patent Nos. 5,834,200 and 5,691,153; for instance, the inventive genes or genes identified by inventive methods herein can be used in methods analogous to those of U.S. Patent Nos. 5,834,200 and 5,691,153.

In the inventive methods steps (a) and/or (b) and optionally (d) and/or (e) are carried out *in vivo* and/or steps (a) and/or (b) and optionally (d) and/or (e) are carried out *in vitro*.

The determination in step (a) and optionally in step (d) can be effected by employing

- (i) a nucleic acid sequence corresponding to at least a part of the gene encoding at least part of the protein and optionally a second nucleic acid sequence corresponding to at least a part of the second gene encoding at least part of the second protein;
- (ii) a nucleic acid sequence complementary to the nucleic acid sequence(s) of (i);
- or
- (iii) a primer or a primer pair hybridizing to the nucleic acid sequence(s) of (i) or (ii).



(Note also the discussion herein, e.g., *infra*, concerning primers/probes and PCR and hybridization.)

The determination in step (b) and optionally of step (c) can be effected by employing an antibody or a fragment thereof that specifically binds to the protein and optionally by employing a second antibody or a fragment thereof which specifically binds to the second protein. (Note also the discussion herein, e.g., *infra*, concerning antibodies and methods for making and uses thereof.)

The medicament or treatment can be any conventional medicament or treatment for osteoporosis. Alternatively or additionally, the medicament or treatment can be the particular protein of the gene detected in the inventive methods, or that which inhibits that protein, e.g., binds to it. Similarly, additionally or alternatively, the medicament or treatment can be a vector which expresses the protein of the gene detected in the inventive methods or that which inhibits expression of that gene; again, for instance, that which can bind to it and/or otherwise prevents its transcription or translation. The selection of administering a protein or that which expresses it, or of administering that which inhibits the protein or the gene expression, can be done without undue experimentation, e.g., based on down regulation or up regulation as determined by inventive methods (e.g., in the osteoporosis model).

In the practice of the invention, one can employ general methods in molecular biology: Standard molecular biology techniques known in the art and not specifically described are generally followed as in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York (1989, 1992), and in Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Maryland (1989).

To determine the absence or decrease from normal cells or presence or increase from normal cells of a nucleic acid molecule, or to amplify it, e.g., in using probes or primers described herein or derived from nucleic acid molecules disclosed herein, the polymerase chain reaction (PCR) may be used and is conveniently carried out generally as in *PCR Protocols: A Guide To Methods And Applications*, Academic Press, San Diego, CA (1990). Reactions and manipulations involving other nucleic acid techniques, unless stated otherwise, are performed as generally described in Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, and methodology as set forth in United States patents 4,666,828;

4,683,202; 4,801,531; 5,192,659 and 5,272,057 and incorporated herein by reference. *In-situ* (In-cell) PCR in combination with Flow Cytometry can be used for detection of cells containing specific DNA and mRNA sequences (Testoni et al, 1996, Blood 87:3822.)

- 5 In PCR, as well as in hybridization, it is preferred that the primers (or probes) bind specifically to the gene of interest, e.g., an inventive gene disclosed herein such as 608, 405 or 274, or a gene identified by methods disclosed herein, or a corresponding human homolog being detected by primer(s) or probe(s) derived from a herein defined gene. One way to ensure this is to select primers from the gene
- 10 sequence that are not generally found in other known sequences.

- The invention accordingly in yet a further aspect provides an isolated nucleic acid molecule, e.g., DNA comprising a sequence encoding a herein defined gene or encoding a herein defined a polypeptide (e.g., an expression product of a herein defined gene) comprising at least about 12 nucleotides in length, for instance, at least
- 15 about 15, about 18, about 21, about 24 or about 27 nucleotides in length, such as at least about 30, about 33, about 36, about 39 or about 42 nucleotides in length, for example, a nucleic acid molecule of at least about 12 nucleotides in length such as about 12 to about 30, about 12 to about 50 or about 12 to about 60, or about 12 to about 75 or about 12 to about 100 or more nucleotides in length. Nucleic acid
- 20 molecules of these lengths may be useful in hybridization; and, the invention further comprehends vectors or plasmids containing and/or expressing such a nucleic acid molecule, as well as uses of such nucleic acid molecules, e.g., for expression thereof either *in vitro* or *in vivo*, or for amplifying or detecting a herein defined gene or a homolog thereof, e.g., a human homolog, in a sample, for instance by a polymerase
- 25 chain reaction.

- A probe or primer can be any stretch of at least 8, preferably at least 10, more preferably at least 12, 13, 14, or 15, such as at least 20, e.g., at least 23 or 25, for instance at least 27 or 30 nucleotides in a herein defined gene which are unique thereto. As to PCR or hybridization primers or probes and optimal lengths therefor,
- 30 reference is also made to Kajimura et al., GATA 7(4):71-79 (1990). The invention will thus be understood to provide oligonucleotides, such as, pairs of oligonucleotides, for use as primers for the *in vitro* amplification of DNA samples and fragments thereof, or for use in expressing a portion of DNA, either *in vitro* or *in vivo*.

The oligonucleotides preferably specifically hybridize to sequences flanking a nucleic acid to be amplified, wherein the oligonucleotides hybridize to different and opposite strands of the double-stranded DNA target. The oligonucleotides of the invention are preferably derived from the nucleic acid molecules, e.g., a herein defined gene, and teachings herein. As used in the practice of this invention, the term "derived from" is intended to encompass the development of such oligonucleotides from the nucleic acid molecules and herein defined gene(s) and teachings disclosed herein, from which a multiplicity of alternative and variant oligonucleotides can be prepared.

The term "specific hybridization" will be understood to mean that the nucleic acid probes of the invention are capable of stable, double-stranded hybridization to gene-derived DNA or RNA under conditions of high stringency, as the term "high stringency" would be understood by those with skill in the art (see, for example, Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. and Hames and Higgins, eds., 1985, Nucleic Acid Hybridization, IRL Press, Oxford, U.K.). Hybridization will be understood to be accomplished using well-established techniques, including but not limited to Southern blot hybridization, Northern blot hybridization, *in situ* hybridization and, most preferably, Southern hybridization to PCR-amplified DNA fragments.

The nucleic acid hybridization probe of the invention may be obtained by use of the polymerase chain reaction (PCR) procedure, using appropriate pairs of PCR oligonucleotide primers as provided herein or derived from the gene sequence(s) provided herein. See U.S. Pat. Nos. 4,683,195 to Mullis et al. and 4,683,202 to Mullis. The invention in a still further aspect provides oligonucleotides for *in vitro* amplification using any of a variety of amplification protocols known in the art. Preferably, the invention provides oligonucleotides for performing polymerase chain reaction (PCR). See U.S. Pat. Nos. 4,683,195 to Mullis et al. and 4,683,202 to Mullis.

The invention will thus be understood to provide oligonucleotides, specifically, pairs of oligonucleotides, for use as primers for the *in vitro* amplification of genes as disclosed herein, e.g., of DNA samples and fragments thereof. In the practice of this invention, the pairs of oligonucleotides herein provided will be understood to comprise two oligonucleotides, comprising from about 8 to about 30 nucleotide residues apiece, said oligonucleotides specifically hybridizing to sequences

- flanking a nucleic acid to be amplified, wherein the oligonucleotides hybridize to different and opposite strands of the DNA target. The oligonucleotides of the invention are preferably derived from the nucleic acid primers discussed below or from the gene(s) disclosed herein. As used in the practice of this invention, the term
- 5 "derived from" is intended to encompass the development of such oligonucleotides from the nucleic acid sequence of the gene(s) or the primers herein disclosed, from which a multiplicity of alternative and variant oligonucleotides can be prepared. In particular, the invention provides oligonucleotides having a sequence that is substantially complementary to the corresponding sequence of the nucleic acid
- 10 hybridization probe. As used herein, the term "substantially corresponding to" is intended to encompass oligonucleotides comprising sequence additions, deletions and mismatches, wherein certain nucleotide residues of the oligonucleotide sequence are not optimally complementary (e.g., A-C or G-T) or are non-complementary (e.g., A-G or T-C) to the corresponding sequence of the nucleic acid hybridization probe,
- 15 provided that such oligonucleotides retain the capacity to specifically amplify the gene(s).

Nucleic acids, e.g., 405, 608 or 274, and oligonucleotides therefrom, such as primers disclosed herein and derivable from the sequences of the present invention (e.g., portions of each disclosed gene which are about 8 to 30 or more nucleotides in

20 length and bind with sufficient specificity to the gene are useful as diagnostic tools for detecting the existence of a osteoporosis or conditions or factors of osteoporosis. Such diagnostic or detection reagents comprise nucleic acid hybridization probes of the invention and encompass paired oligonucleotide PCR primers, as described above.

Methods provided by the invention include blot hybridization, *in situ*

25 hybridization and *in vitro* amplification techniques for detecting osteoporosis or conditions or factors of osteoporosis in a sample such as a biological sample. Appropriate biological samples advantageously screened using the methods described herein include blood, serum, saliva and other body fluids, and other potential sources of infection.

30 In the detection methods of the invention, production of a specific DNA fragment produced by *in vitro* amplification of a template DNA sample is detected by agarose gel electrophoresis, ethidium bromide staining and ultraviolet transillumination of ethidium bromide stained gels, performed using conventional

techniques (Sambrook et al., *supra*), or detection by sequence detection systems using fluorogenic or other labeled probes that rely on automatic or automated detection instrumentation. In instances where a greater degree of specificity is required, hybridization of such agarose gels probed with a detectably-labeled nucleic acid hybridization probe of the invention is performed using standard techniques (Sambrook et al., *supra*). In each of these embodiments of the methods of the invention, a sufficient amount of a specific PCR-amplified DNA fragment is produced to be readily detected. For the purposes of this invention, the term "a sufficient amount of a specific PCR-amplified DNA fragment" is defined as that amount required to be detected, either by visualization of ethidium bromide-stained agarose gels or autoradiographic or other development of a blot hybridized with a detectably-labeled probe.

It will be understood that a sufficient quantity of a specific PCR amplified DNA fragment is prepared in PCR amplification reactions by performing a number of cycles required to produce said sufficient amount of the specific DNA fragment. The number of cycles in each PCR required to produce said sufficient amount of a specific DNA fragment will be understood to depend on the oligonucleotide primers, buffers, salts and other reaction components, the amount of template DNA and the PCR cycling times and temperatures. It will also be understood that the optimization of these parameters are within the skill of the worker of ordinary skill to achieve with no more than routine experimentation.

Detectably-labeled probes as provided by the invention are labeled with biotin, a radioisotope (including  $^3\text{H}$ ,  $^{14}\text{C}$ ,  $^{32}\text{S}$  and  $^{32}\text{P}$ ), a fluorescent label (including fluorescein isothiocyanate), and an antigenic label. The detectable label is incorporated into the probe during synthetic preparation of the probe, whereby the probe is alternatively end-labeled or labeled by the incorporation of labeled nucleotides into the synthesized probe.

The invention also provides a PCR-based method for preparing a nucleic acid hybridization probe of the invention. In these embodiments, template DNA comprises a recombinant genetic construct of the invention. A detectably-labeled nucleic acid hybridization probe is prepared by performing PCR amplification using a pair of oligonucleotide primers specific for sequences flanking the position of the nucleic

acid insert. Detectable label is incorporated into the nucleic acid hybridization probe by direct end-labeling of PCR primers or incorporation of detectably-labeled nucleotide triphosphates into the probe nucleic acid.

PCR comprising the methods of the invention is performed in a reaction mixture comprising an amount, typically between <10 ng-200 ng template nucleic acid; 50-100 pmoles each oligonucleotide primer; 1-1.25 mM each deoxynucleotide triphosphate; a buffer solution appropriate for the polymerase used to catalyze the amplification reaction; and 0.5-2 Units of a polymerase, most preferably a thermostable polymerase (e.g., Taq polymerase or Tth polymerase).

The invention thus provides in further aspects diagnostic assays for the specific detection of osteoporosis or genes associated therewith. These diagnostic assays include nucleic acid hybridization assays, using the nucleic acids of the invention or specifically-hybridizing fragments thereof, for sensitive detection of fungal genomic DNA and/or RNA. Such assays include various blot assays, such as Southern blots, Northern blots, dot blots, slot blots and the like, as well as *in vitro* amplification assays, such as the polymerase chain reaction assay (PCR), reverse transcription-polymerase chain reaction assay (RT-PCR), ligase chain reaction assay (LCR), and others known to those skilled in the art. Specific restriction endonuclease digestion of diagnostic fragments detected using any of the methods of the invention, analogous to restriction fragment linked polymorphism assays (RFLP) are also within the scope of this invention.

These PCR techniques can be used in conjunction with or in the practice of other methods disclosed herein, or other conditions associated with or correlated to mechanical stress or a lack thereof.

Accordingly, the invention relates to compositions and methods for detecting and/or diagnosing osteoporosis or conditions or factors associated therewith, including genetic factors associated therewith.

Similarly, in the practice of the invention, e.g., protein detection, general methods in immunology may be employed. Standard methods in immunology known in the art and not specifically described are generally followed as in Stites et al. (eds), Basic and Clinical Immunology (8th Edition), Appleton & Lange, Norwalk, CT (1994) and Mishell and Shiigi (eds), Selected Methods in Cellular Immunology, W.H. Freeman and Co., New York (1980). Immunoassays such as RIA and ELISA can be

employed to assess a specimen for the presence of specific proteins or other compounds of interest where appropriate as known in the art. Both polyclonal and monoclonal antibodies can be used in the assays. Available immunoassays are extensively described in the patent and scientific literature. See, for example, United States patents 3,791,932; 3,839,153; 3,850,752; 3,850,578; 3,853,987; 3,867,517; 3,879,262; 3,901,654; 3,935,074; 3,984,533; 3,996,345; 4,034,074; 4,098,876; 4,879,219; 5,011,771 and 5,281,521 as well as Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Cold Springs Harbor, New York, 1989

- Antibodies may be used in various aspects of the invention, e.g., in detection or treatment or prevention methods. Antibodies may be either monoclonal, polyclonal or recombinant to be used in the immunoassays or other methods of analysis necessary for the practice of the invention. Conveniently, the antibodies may be prepared against the immunogen or portion thereof for example a synthetic peptide based on the sequence, or prepared recombinantly by cloning techniques or the natural gene product and/or portions thereof may be isolated and used as the immunogen. The genes are identified as set forth in the present invention and the gene product identified. Immunogens can be used to produce antibodies by standard antibody production technology well known to those skilled in the art as described generally in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988 and Borrebaeck, *Antibody Engineering - A Practical Guide*, W.H. Freeman and Co., 1992. Antibody fragments may also be prepared from the antibodies and include Fab, F(ab')<sub>2</sub>, and Fv by methods known to those skilled in the art.

- For producing polyclonal antibodies a host, such as a rabbit or goat, is immunized with the immunogen or immunogen fragment, generally with an adjuvant and, if necessary, coupled to a carrier; antibodies to the immunogen are collected from the sera. Further, the polyclonal antibody can be absorbed such that it is monospecific. That is, the sera can be absorbed against related immunogens so that no cross-reactive antibodies remain in the sera rendering it monospecific.

- For producing monoclonal antibodies the technique involves hyperimmunization of an appropriate donor with the immunogen, generally a mouse, and isolation of splenic antibody producing cells. These cells are fused to a cell having immortality, such as a myeloma cell, to provide a fused cell hybrid which has

immortality and secretes the required antibody. The cells are then cultured, in bulk, and the monoclonal antibodies harvested from the culture media for use.

For producing recombinant antibody (see generally Huston et al, 1991; Johnson and Bird, 1991; Mernaugh and Mernaugh, 1995), messenger RNAs from antibody producing Blymphocytes of animals, or hybridoma are reverse -transcribed to obtain complimentary DNAs (cDNAs). Antibody cDNA, which can be full or partial length, is amplified and cloned into a phage or a plasmid. The cDNA can be a partial length of heavy and light chain cDNA, separated or connected by a linker. The antibody, or antibody fragment, is expressed using a suitable expression system to obtain recombinant antibody. Antibody cDNA can also be obtained by screening pertinent expression libraries.

The antibody can be bound to a solid support substrate or conjugated with a detectable moiety or be both bound and conjugated as is well known in the art. (For a general discussion of conjugation of fluorescent or enzymatic moieties see Johnstone & Thorpe, *Immunochemistry in Practice*, Blackwell Scientific Publications, Oxford, 1982.) The binding of antibodies to a solid support substrate is also well known in the art. (see for a general discussion Harlow & Lane *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Publications, New York, 1988 and Borrebaeck, *Antibody Engineering - A Practical Guide*, W.H. Freeman and Co., 1992) The detectable moieties contemplated with the present invention can include, but are not limited to, fluorescent, metallic, enzymatic and radioactive markers such as biotin, gold, ferritin, alkaline phosphatase, Beta-galactosidase, peroxidase, urease, fluorescein, rhodamine, tritium,  $^{13}\text{C}$  and iodination.

Antibodies can also be used as an active agent in a therapeutic composition and such antibodies can be humanized, for instance, to enhance their effects. See, e.g., Huls et al., "A recombinant, fully human monoclonal antibody with antitumor activity constructed from phage-displayed antibody fragments," *Nature Biotechnology* Vol. 17, No. 3, March 1999, and documents cited therein, incorporated herein by reference.

Accordingly, antibodies from expression products of genes identified herein or by inventive methods disclosed herein are useful in immunodiagnostics, as well as in drugs or other commercial uses such as in research.



Simply, the expression product from the gene or portions thereof can be useful for generating antibodies such as monoclonal or polyclonal antibodies which are useful for diagnostic purposes or to block activity of expression products or portions thereof or of genes or a portion thereof, e.g., as a therapeutic. Monoclonal antibodies are immunoglobulins produced by hybridoma cells. A monoclonal antibody reacts with a single antigenic determinant and provides greater specificity than a conventional, serum-derived antibody. Furthermore, screening a large number of monoclonal antibodies makes it possible to select an individual antibody with desired specificity, avidity and isotype. Hybridoma cell lines provide a constant, inexpensive source of chemically identical antibodies and preparations of such antibodies can be easily standardized. Methods for producing monoclonal antibodies are well known to those of ordinary skill in the art, e.g., Koprowski, H. et al., U.S. Pat. No. 4,196,265, issued Apr. 1, 1989, incorporated herein by reference, and other documents cited herein, e.g., *supra*.

Uses of monoclonal antibodies are known. One such use is in David, G. and Greene, H., U.S. Pat. No. 4,376,110, issued Mar. 8, 1983, incorporated herein by reference; *see also* documents cited herein, e.g., *supra*. Monoclonal antibodies have also been used to recover materials by immunoadsorption chromatography, *see, e.g.* Milstein, C., 1980, Scientific American 243:66, 70, incorporated herein by reference; and documents cited herein, such as *supra*. Thus, products expressed from genes identified herein or by methods herein or portions thereof are useful in therapeutics, immunoadsorption chromatography, as well as for generating antibodies for diagnostic or detection purposes. Furthermore, the expression products can be used in assays for detecting the presence of antibodies. For instance, the antibodies or expressed products can be used in assays analogous to those disclosed in U.S. Patents Nos. 5,591,645, 4,861,711, 5,861,319, 5,858,804, and 5,863,720, as well as in WO 86/04683, EP 154 749, WO 86/03839, and EP 186 799. Antibodies in the practice of the invention can include fragments thereof which are functional, e.g., a fragment that at least statistically significantly retains some (for instance a majority) or all of binding as compared with the entire antibody; for instance, antibodies comprehends a fragment comprising a binding domain.

Protein purification, including recombinant protein purification in the practice of the invention can be in accordance with or analogous to Marshak et al, "Strategies

for Protein Purification and Characterization. A laboratory course manual. " CSHL Press, 1996.

With respect to transgenic and knockout methods, the present invention comprehends transgenic gene and polymorphic gene animal and cellular (cell lines) models as well as for knockout models for the genes identified in the present invention. These models are constructed using standard methods known in the art and as set forth in United States Patents 5,487,992, 5,464,764, 5,387,742, 5,360,735, 5,347,075, 5,298,422, 5,288,846, 5,221,778, 5,175,385, 5,175,384, 5,175,383, 4,736,866 as well as Burke and Olson (1991), Capecchi (1989), Davies et al. (1992), Dickinson et al. (1993), Duff and Lincoln (1995), Huxley et al. (1991), Jakobovits et al. (1993), Lamb et al. (1993), Pearson and Choi (1993), Rothstein (1991), Schedl et al. (1993), Strauss et al. (1993). Further, patent applications WO 94/23049, WO 93/14200, WO 94/06908, WO 94/28123 also provide information.

Thus, for instance, the inventive methods of the invention can be used to determine a gene of interest with respect to a physiological or disease state, e.g., osteoporosis or other conditions caused by mechanical stress for instance 608 or 405 or 274 and using the information herein and in the art (such as documents cited herein), knockout or transgenic animals such as mice or rat, can be prepared, to generate animals prone to the physiological or disease state, osteoporosis or other conditions caused by mechanical stress, to thereby test treatments or medicaments therefor, or, to test theories and thus advance research pertaining to the physiological or disease state, e.g., to test the functions of identified genes such as 405, 608 and 274, *inter alia*. Accordingly, from this disclosure and the knowledge in the art, no undue experimentation is needed to prepare knockout or transgenic animals, such as mice or rats or rodents; and, such animals have great value and utility.

Moreover, the genes of the present invention or a portion thereof, e.g., a portion thereof which expresses a protein which function the same as or analogously to the full length protein, or genes identified by the methods herein can be expressed recombinantly, e.g., in *E. coli* or in another vector or plasmid for either *in vivo* expression or *in vitro* expression. The methods for making and/or administering a vector or recombinant or plasmid for expression of gene products of genes of the invention or identified by the invention or a portion thereof either *in vivo* or *in vitro* can be any desired method, e.g., a method which is by or analogous to the methods

- disclosed in: U.S. Patent Nos. 4,603,112, 4,769,330, 5,174,993, 5,505,941, 5,338,683, 5,494,807, 4,722,848, WO 94/16716, WO 96/39491, Paoletti, "Applications of pox virus vectors to vaccination: An update," PNAS USA 93:11349-11353, October 1996, Moss, "Genetically engineered poxviruses for recombinant gene expression, vaccination, and safety," PNAS USA 93:11341-11348, October 1996, Smith et al., U.S. Patent No. 4,745,051 (recombinant baculovirus), Richardson, C.D. (Editor), Methods in Molecular Biology 39, "Baculovirus Expression Protocols" (1995 Humana Press Inc.), Smith et al., "Production of Human Beta Interferon in Insect Cells Infected with a Baculovirus Expression Vector," Molecular and Cellular Biology, Dec., 1983, Vol. 3, No. 12, p. 2156-2165; Pennock et al., "Strong and Regulated Expression of *Escherichia coli* B-Galactosidase in Infect Cells with a Baculovirus vector," Molecular and Cellular Biology Mar. 1984, Vol. 4, No. 3, p. 399-406; EPA 0 370 573, U.S. application Serial No. 920,197, filed October 16, 1986, EP Patent publication No. 265785, U.S. Patent No. 4,769,331 (recombinant herpesvirus), Roizman, "The function of herpes simplex virus genes: A primer for genetic engineering of novel vectors," PNAS USA 93:11307-11312, October 1996, Andreansky et al., "The application of genetically engineered herpes simplex viruses to the treatment of experimental brain tumors," PNAS USA 93:11313-11318, October 1996, Robertson et al. "Epstein-Barr virus vectors for gene delivery to B lymphocytes," PNAS USA 93:11334-11340, October 1996, Frolov et al., "Alphavirus-based expression vectors: Strategies and applications," PNAS USA 93:11371-11377, October 1996, Kitson et al., J. Virol. 65, 3068-3075, 1991; U.S. Patent Nos. 5,591,439, 5,552,143 (recombinant adenovirus), Grunhaus et al., 1992, "Adenovirus as cloning vectors," Seminars in Virology (Vol. 3) p. 237-52, 1993, Ballay et al. EMBO Journal, vol. 4, p. 3861-65, Graham, Tibtech 8, 85-87, April, 1990, Prevec et al., J. Gen Virol. 70, 429-434, PCT WO91/11525, Felgner et al. (1994), J. Biol. Chem. 269, 2550-2561, Science, 259:1745-49, 1993 and McClements et al., "Immunization with DNA vaccines encoding glycoprotein D or glycoprotein B, alone or in combination, induces protective immunity in animal models of herpes simplex virus-2 disease," PNAS USA 93:11414-11420, October 1996, and U.S. Patents Nos 5,591,639, 5,589,466, and 5,580,859 relating to DNA expression vectors, *inter alia*. See also WO 98/33510; Ju et al., Diabetologia, 41:736-739, 1998 (lentiviral expression system); Sanford et al., U.S. Patent No. 4,945,050 (method for

transporting substances into living cells and tissues and apparatus therefor); Fischbach et al. (Intracel), WO 90/01543 (method for the genetic expression of heterologous proteins by cells transfected); Robinson et al., seminars in IMMUNOLOGY, vol. 9, pp.271-283 (1997) (DNA vaccines); Szoka et al., U.S. Patent No. 4,394,448 (method of inserting DNA into living cells); and McCormick et al., U.S. Patent No. 5,677,178 (use of cytopathic viruses for therapy and prophylaxis of neoplasia).

The expression product generated by vectors or recombinants in this invention optionally can also be isolated and/or purified from infected or transfected cells; for instance, to prepare compositions for administration to patients. However, in certain instances, it may be advantageous to not isolate and/or purify an expression product from a cell; for instance, when the cell or portions thereof enhance the effect of the polypeptide.

An inventive vector or recombinant expressing a gene identified herein or from a method herein or a portion thereof can be administered in any suitable amount to achieve expression at a suitable dosage level, e.g., a dosage level analogous to the herein mentioned dosage levels (wherein the gene product is directly present). The inventive vector or recombinant can be administered to a patient or infected or transfected into cells in an amount of about at least  $10^3$  pfu; more preferably about  $10^4$  pfu to about  $10^{10}$  pfu, e.g., about  $10^5$  pfu to about  $10^9$  pfu, for instance about  $10^6$  pfu to about  $10^8$  pfu. In plasmid compositions, the dosage should be a sufficient amount of plasmid to elicit a response analogous to compositions wherein gene product or a portion thereof is directly present; or to have expression analogous to dosages in such compositions; or to have expression analogous to expression obtained *in vivo* by recombinant compositions. For instance, suitable quantities of plasmid DNA in plasmid compositions can be 1  $\mu$ g to 100 mg, preferably 0.1 to 10 mg, e.g., 500 micrograms, but lower levels such as 0.1 to 2 mg or preferably 1-10  $\mu$ g may be employed. Documents cited herein regarding DNA plasmid vectors may be consulted for the skilled artisan to ascertain other suitable dosages for DNA plasmid vector compositions of the invention, without undue experimentation.

Compositions for administering vectors can be as in or analogous to such compositions in documents cited herein or as in or analogous to compositions herein described, e.g., pharmaceutical or therapeutic compositions and the like (e.g., *see infra*).

Thus, the invention comprehends *in vivo* gene expression which is sometimes termed "gene therapy". Gene therapy can refer to the transfer of genetic material (e.g. DNA or RNA) of interest into a host to treat or prevent a genetic or acquired disease or condition phenotype. The particular gene that is to be used or which has been identified as the target gene is identified as set forth herein. The genetic material of interest encodes a product (e.g. a protein, polypeptide, peptide or functional RNA) whose production *in vivo* is desired. For example, the genetic material of interest can encode a hormone, receptor, enzyme, polypeptide or peptide of therapeutic value. For a review see, in general, the text "Gene Therapy" (Advances in Pharmacology 40, Academic Press, 1997).

Two basic approaches to gene therapy have evolved: (1) *ex vivo* and (2) *in vivo* gene therapy. In *ex vivo* gene therapy cells are removed from a patient, and while being cultured are treated *in vitro*. Generally, a functional replacement gene is introduced into the cell via an appropriate gene delivery vehicle/method (transfection, homologous recombination, etc.) and, an expression system as needed and then the modified cells are expanded in culture and returned to the host/patient. These genetically reimplanted cells have been shown to produce the transfected gene product *in situ*. In *in vivo* gene therapy, target cells are not removed from the subject rather the gene to be transferred is introduced into the cells of the recipient organism *in situ*, that is within the recipient. Alternatively, if the host gene is defective, the gene is repaired *in situ* (Culver, 1998). These genetically altered cells have been shown to produce the transfected gene product *in situ*.

The gene expression vehicle is capable of delivery/transfer of heterologous nucleic acid into a host cell. The expression vehicle may include elements to control targeting, expression and transcription of the nucleic acid in a cell selective manner as is known in the art. It should be noted that often the 5'UTR and/or 3'UTR of the gene may be replaced by the 5' UTR and/or 3'UTR of the expression vehicle. Therefore as used herein the expression vehicle may, as needed, not include the 5'UTR and/or 3'UTR shown in sequences herein and only include the specific amino acid coding region.

The expression vehicle can include a promotor for controlling transcription of the heterologous material and can be either a constitutive or inducible promotor to allow selective transcription. Enhancers that may be required to obtain necessary

transcription levels can optionally be included. Enhancers are generally any non-translated DNA sequence which works contiguously with the coding sequence (in cis) to change the basal transcription level dictated by the promoter. The expression vehicle can also include a selection gene as described herein.

- 5        Vectors can be introduced into cells or tissues by any one of a variety of known methods within the art. Such methods can be found generally described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Springs Harbor Laboratory, New York (1989, 1992), in Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Maryland (1989), Chang et al., *Somatic*  
10 *Gene Therapy*, CRC Press, Ann Arbor, MI (1995), Vega et al., *Gene Targeting*, CRC Press, Ann Arbor, MI (1995), *Vectors: A Survey of Molecular Cloning Vectors and Their Uses*, Butterworths, Boston MA (1988) and Gilboa et al (1986), as well as other documents cited herein (*see supra*) and include, for example, stable or transient transfection, lipofection, electroporation and infection with recombinant viral vectors.
- 15 In addition, see United States patent 4,866,042 for vectors involving the central nervous system and also United States patents 5,464,764 and 5,487,992 for positive-negative selection methods.

- Introduction of nucleic acids by infection offers several advantages over the other listed methods. Higher efficiency can be obtained due to their infectious nature.
- 20 Moreover, viruses are very specialized and typically infect and propagate in specific cell types. Thus, their natural specificity can be used to target the vectors to specific cell types in vivo or within a tissue or mixed culture of cells. Viral vectors can also be modified with specific receptors or ligands to alter target specificity through receptor mediated events.

- 25        A specific example of DNA viral vector for introducing and expressing recombinant sequences is the adenovirus derived vector Adenop53TK. This vector expresses a herpes virus thymidine kinase (TK) gene for either positive or negative selection and an expression cassette for desired recombinant sequences. This vector can be used to infect cells that have an adenovirus receptor which includes most
- 30 cancers of epithelial origin as well as others. This vector as well as others that exhibit similar desired functions, can be used to treat a mixed population of cells and can include, for example, an *in vitro* or *ex vivo* culture of cells, a tissue or a human subject.

Additional features can be added to the vector to ensure its safety and/or enhance its therapeutic efficacy. Such features include, for example, markers that can be used to negatively select against cells infected with the recombinant virus. An example of such a negative selection marker is the TK gene described above that confers sensitivity to the antibiotic gancyclovir. Negative selection is therefore a means by which infection can be controlled because it provides inducible suicide through the addition of antibiotic. Such protection ensures that if, for example, mutations arise that produce altered forms of the viral vector or recombinant sequence, cellular transformation will not occur. Features that limit expression to particular cell types can also be included. Such features include, for example, promoter and regulatory elements that are specific for the desired cell type.

In addition, recombinant viral vectors are useful for *in vivo* expression of a desired nucleic acid because they offer advantages such as lateral infection and targeting specificity. Lateral infection is inherent in the life cycle of, for example, retrovirus and is the process by which a single infected cell produces many progeny virions that bud off and infect neighboring cells. The result is that a large area becomes rapidly infected, most of which was not initially infected by the original viral particles. This is in contrast to vertical-type of infection in which the infectious agent spreads only through daughter progeny. Viral vectors can also be produced that are unable to spread laterally. This characteristic can be useful if the desired purpose is to introduce a specified gene into only a localized number of targeted cells.

As described above, viruses are very specialized infectious agents that have evolved, in many cases, to elude host defense mechanisms. Typically, viruses infect and propagate in specific cell types. The targeting specificity of viral vectors utilizes its natural specificity to specifically target predetermined cell types and thereby introduce a recombinant gene into the infected cell. The vector to be used in the methods of the invention will depend on desired cell type to be targeted and will be known to those skilled in the art. For example, if breast cancer is to be treated then a vector specific for such epithelial cells would be used. Likewise, if diseases or pathological conditions of the hematopoietic system are to be treated, then a viral vector that is specific for blood cells and their precursors, preferably for the specific type of hematopoietic cell, would be used.

Retroviral vectors can be constructed to function either as infectious particles or to undergo only a single initial round of infection. In the former case, the genome of the virus is modified so that it maintains all the necessary genes, regulatory sequences and packaging signals to synthesize new viral proteins and RNA. Once these molecules are synthesized, the host cell packages the RNA into new viral particles which are capable of undergoing further rounds of infection. The vector's genome is also engineered to encode and express the desired recombinant gene. In the case of non-infectious viral vectors, the vector genome is usually mutated to destroy the viral packaging signal that is required to encapsulate the RNA into viral particles. Without such a signal, any particles that are formed will not contain a genome and therefore cannot proceed through subsequent rounds of infection. The specific type of vector will depend upon the intended application. The actual vectors are also known and readily available within the art or can be constructed by one skilled in the art using well-known methodology.

The recombinant vector can be administered in several ways. If viral vectors are used, for example, the procedure can take advantage of their target specificity and consequently, do not have to be administered locally at the diseased site. However, local administration can provide a quicker and more effective treatment, administration can also be performed by, for example, intravenous or subcutaneous injection into the subject. Injection of the viral vectors into a spinal fluid can also be used as a mode of administration, especially in the case of neurodegenerative diseases. Following injection, the viral vectors will circulate until they recognize host cells with the appropriate target specificity for infection.

An alternate mode of administration can be by direct inoculation locally at the site of the disease or pathological condition or by inoculation into the vascular system supplying the site with nutrients or into the spinal fluid. Local administration is advantageous because there is no dilution effect and, therefore, a smaller dose is required to achieve expression in a majority of the targeted cells. Additionally, local inoculation can alleviate the targeting requirement required with other forms of administration since a vector can be used that infects all cells in the inoculated area. If expression is desired in only a specific subset of cells within the inoculated area, then promoter and regulatory elements that are specific for the desired subset can be used to accomplish this goal. Such non-targeting vectors can be, for example, viral vectors,



viral genome, plasmids, phagemids and the like. Transfection vehicles such as liposomes can also be used to introduce the non-viral vectors described above into recipient cells within the inoculated area. Such transfection vehicles are known by one skilled within the art.

- 5           Inventive vectors can comprise a herein defined gene, as well as a regulatory element operative linked thereto, e.g., a promoter, for expression; and, the regulatory element or promoter can be tissue or cell specific; for instance, the regulatory element or promoter can be for expression in a cell or precursor thereto employed in an inventive or herein-described or herein-cited test, e.g., the regulatory element or
- 10          promoter can be for expression in a bone cell such as an osteoblast or an osteoclast or a precursor thereto.

- Delivery of gene products (products from herein defined genes: genes identified herein or by inventive methods or portions thereof) and/or antibodies or portions thereof and/or agonists or antagonists (collectively or individually
- 15          "therapeutics"), and compositions comprising the same, as well as of compositions comprising a vector expressing gene products, can be done without undue experimentation from this disclosure and the knowledge in the art.

- The present invention provides compositions comprising an expression vector comprising a herein defined gene, or a portion thereof, e.g., which codes for a
- 20          functional portion thereof, as well as therapeutics based on the genes identified herein, e.g., compositions comprising expression products or a functional portion thereof or antibodies thereto or a functional portion thereof and/or agonists or antagonists.
- (Thus, a herein defined gene can comprehend a portion thereof which expresses a functional portion of a full length expression product.). The therapeutics and vectors
- 25          of the present invention are administered and dosed in accordance with good medical practice, taking into account the clinical condition of the individual patient, the site and method of administration, scheduling of administration, patient age, sex, body weight, species of the patient, and other factors known to those skilled in the pharmaceutical or veterinary arts.
- 30          The pharmaceutically "effective amount" for purposes herein is thus determined by such considerations as are known in the art. The amount must be effective to achieve improvement including but not limited to improved survival rate or more rapid recovery, or improvement or elimination of symptoms and other

indicators, e.g., of osteoporosis, for instance, improvement in bone density, as are selected as appropriate measures by those skilled in the art.

Where appropriate the therapeutics of the present invention are pharmaceuticals and as such can be administered in various ways. It should be noted that these therapeutics can be administered as the expression product and/or portion thereof and/or antibody and/or portion thereof or as pharmaceutically acceptable salt and can be administered alone or as an active ingredient in combination with pharmaceutically acceptable carriers, diluents, adjuvants and vehicles, as well as other active ingredients (e.g., other expression products, portions thereof, antibodies, portions thereof, from inventive methods, and/or other therapies, such as those discussed herein). The compounds can be administered orally, subcutaneously or parenterally including intravenous, intraarterial, intramuscular, intraperitoneally, and intranasal administration as well as intrathecal and infusion techniques.

Implants of the therapeutics and/or of vectors expressing the herein defined genes are also useful. The patient being treated is a warm-blooded animal and, in particular, mammals including man. The pharmaceutically acceptable carriers, diluents, adjuvants and vehicles as well as implant carriers generally refer to inert, non-toxic solid or liquid fillers, diluents or encapsulating material not reacting with the active ingredients of the invention. The implant can be placed near bone, to stimulate bone growth or increase bone density. With respect to implants or slow release systems that can be used in the practice of the invention with respect to therapeutics, or vectors expressing the herein defined genes, mention is made of U.S. Patents Nos. 4,150,108, 4,329,332, 4,331,652, 4,333,919, 4,389,330, 4,489,055, 4,526,938, 4,530,840, 4,542,025, 4,563,489, 4,675,189, 4,677,191, 4,683,288, 4,758,435, 4,857,335, 4,931,287, 5,178,872, 5,252,701, 5,275,820, 5,478,564, 5,540,912, 5,447,725, 5,599,852, 5,607,686, 5,609,886, 5,631,015, 5,654,010, 5,700,485, 5,702,717, 5,711,968, 5,733,566, 4,938,763, 5,077,049, 5,278,201, 5,278,202, 5,288,496, 5,324,519, 5,324,520, 5,340,849, 5,368,859, 5,401,507, 5,419,910, 5,427,796, 5,487,897, 5,599,552, 5,632,727, 5,643,595, 5,660,849, 5,686,092, 5,702,716, 5,707,647, 5,717,030, 5,725,491, 5,733,950, 5,736,152, 5,744,153, 5,759,563, and 5,780,044, European Patent Application 0537559, Shah et al (J. Controlled Release, 1993, 27:139-147), Lambert and Peck (J. Controlled

Release, 1995, 33:189-195), and Shivley et al (J. Controlled Release, 1995, 33:237-243).

It is noted that humans are treated generally longer than the mice or other experimental animals exemplified herein which treatment has a length proportional to the length of the disease process and drug effectiveness. The doses may be single doses or multiple doses over a period of several days, but single doses are preferred. Thus, one can scale up from animal experiments, e.g., rats, mice, and the like, to humans, by techniques from this disclosure and the knowledge in the art, without undue experimentation.

The doses may be single doses or multiple doses over a period of several days. The treatment generally has a length proportional to the length of the disease process and drug effectiveness and the patient species being treated.

When administering a therapeutic or vector of the present invention parenterally, it will generally be formulated in a unit dosage injectable form (solution, suspension, emulsion). The pharmaceutical formulations suitable for injection include sterile aqueous solutions or dispersions and sterile powders for reconstitution into sterile injectable solutions or dispersions. The carrier can be a solvent or dispersing medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils.

Proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Nonaqueous vehicles such as cottonseed oil, sesame oil, olive oil, soybean oil, corn oil, sunflower oil, or peanut oil and esters, such as isopropyl myristate, may also be used as solvent systems for compound compositions

Additionally, various additives which enhance the stability, sterility, and isotonicity of the compositions, including antimicrobial preservatives, antioxidants, chelating agents, and buffers, can be added. Prevention of the action of microorganisms can be ensured by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, and the like. In many cases, it will be desirable to include isotonic agents, for example, sugars, sodium chloride, and the like. Prolonged absorption of the injectable pharmaceutical form can be brought about by the use of agents delaying absorption, for example, aluminum monostearate

and gelatin. According to the present invention, however, any vehicle, diluent, or additive used would have to be compatible with the compounds.

Sterile injectable solutions can be prepared by incorporating the compounds utilized in practicing the present invention in the required amount of the appropriate  
5 solvent with various of the other ingredients, as desired.

A pharmacological formulation of the present invention, e.g., comprising a therapeutic and/or vector, can be administered to the patient in an injectable formulation containing any compatible carrier, such as various vehicle, adjuvants, additives, and diluents; or the compounds utilized in the present invention can be  
10 administered parenterally to the patient in the form of slow-release subcutaneous implants or targeted delivery systems such as monoclonal antibodies, vectored delivery, iontophoretic, polymer matrices, liposomes, and microspheres. Examples of delivery systems useful in the present invention include: 5, 225, 182; 5, 169, 383; 5, 167, 616; 4,959,217; 4,925,678; 4,487,603; 4,486,194; 4,447,233; 4,447,224;  
15 4,439,196; and 4,475,196. Many other such implants, delivery systems, and modules are well known to those skilled in the art (*See also* documents cited herein, e.g., *supra*).

A pharmacological formulation of the compound utilized in the present invention can be administered orally to the patient. Conventional methods such as  
20 administering the compounds in tablets, suspensions, solutions, emulsions, capsules, powders, syrups and the like are usable. Known techniques which deliver it orally or intravenously and retain the biological activity are preferred.

In one embodiment, a formulation of the present invention can be administered initially, and thereafter maintained by further administration. For  
25 instance, a formulation of the invention can be administered in one type of composition and thereafter further administered in a different or the same type of composition. For example, a formulation of the invention can be administered by intravenous injection to bring blood levels to a suitable level. The patient's levels are then maintained by an oral dosage form, although other forms of administration,  
30 dependent upon the patient's condition and as indicated above, can be used.

The quantity to be administered will vary for the patient being treated and will vary from about 100 ng/kg of body weight to 100 mg/kg of body weight per day and preferably will be from 10 pg/kg to 10 mg/kg per day. For instance, dosages can be

readily ascertained by those skilled in the art from this disclosure and the knowledge in the art. Thus, the skilled artisan can readily determine the amount of gene product and optional additives, vehicles, carrier and/or adjuvant in compositions and to be administered in methods of the invention. Typically, an adjuvant or additive is commonly used as 0.001 to 50 wt% solution in phosphate buffered saline, and the gene product or active ingredient is present on the order of micrograms to milligrams, such as about 0.0001 to about 5 wt%, preferably about 0.0001 to about 1 wt%, most preferably about 0.0001 to about 0.05 wt% or about 0.001 to about 20 wt%, preferably about 0.01 to about 10 wt%, and most preferably about 0.05 to about 5 wt%. Of course, for any composition to be administered to an animal or human, including the components thereof, and for any particular method of administration, it is preferred to determine therefor: toxicity, such as by determining the lethal dose (LD) and LD<sub>50</sub> in a suitable animal model e.g., rodent such as mouse; and, the dosage of the composition(s), concentration of components therein and timing of administering the composition(s), which elicit a suitable response, such as by titrations of sera and analysis thereof, e.g., by ELISA and/or RFFIT analysis. Such determinations do not require undue experimentation from the knowledge of the skilled artisan, this disclosure and the documents cited herein. And, the time for sequential administrations can be ascertained without undue experimentation.

Examples of compositions comprising a therapeutic and/or vector of the invention include liquid preparations for orifice, e.g., oral, nasal, anal, vaginal, peroral, intragastric, mucosal (e.g., perlingual, alveolar, gingival, olfactory or respiratory mucosa) etc., administration such as suspensions, syrups or elixirs; and, preparations for parenteral, subcutaneous, intradermal, intramuscular or intravenous administration (e.g., injectable administration), such as sterile suspensions or emulsions. Such compositions may be in admixture with a suitable carrier, diluent, or excipient such as sterile water, physiological saline, glucose or the like. The compositions can also be lyophilized. The compositions can contain auxiliary substances such as wetting or emulsifying agents, pH buffering agents, gelling or viscosity enhancing additives, preservatives, flavoring agents, colors, and the like, depending upon the route of administration and the preparation desired. Standard texts, such as "REMINGTON'S PHARMACEUTICAL SCIENCE", 17th edition,

1985, incorporated herein by reference, may be consulted to prepare suitable preparations, without undue experimentation.

Compositions of the invention, are conveniently provided as liquid preparations, e.g., isotonic aqueous solutions, suspensions, emulsions or viscous compositions which may be buffered to a selected pH. If digestive tract absorption is preferred, compositions of the invention can be in the "solid" form of pills, tablets, capsules, caplets and the like, including "solid" preparations which are time-released or which have a liquid filling, e.g., gelatin covered liquid, whereby the gelatin is dissolved in the stomach for delivery to the gut. If nasal or respiratory (mucosal) administration is desired, compositions may be in a form and dispensed by a squeeze spray dispenser, pump dispenser or aerosol dispenser. Aerosols are usually under pressure by means of a hydrocarbon. Pump dispensers can preferably dispense a metered dose or, a dose having a particular particle size.

Compositions of the invention can contain pharmaceutically acceptable flavors and/or colors for rendering them more appealing, especially if they are administered orally. The viscous compositions may be in the form of gels, lotions, ointments, creams and the like (e.g., for transdermal administration) and will typically contain a sufficient amount of a thickening agent so that the viscosity is from about 2500 to 6500 cps, although more viscous compositions, even up to 10,000 cps may be employed. Viscous compositions have a viscosity preferably of 2500 to 5000 cps, since above that range they become more difficult to administer. However, above that range, the compositions can approach solid or gelatin forms which are then easily administered as a swallowed pill for oral ingestion.

Liquid preparations are normally easier to prepare than gels, other viscous compositions, and solid compositions. Additionally, liquid compositions are somewhat more convenient to administer, especially by injection or orally, to animals, children, particularly small children, and others who may have difficulty swallowing a pill, tablet, capsule or the like, or in multi-dose situations. Viscous compositions, on the other hand, can be formulated within the appropriate viscosity range to provide longer contact periods with mucosa, such as the lining of the stomach or nasal mucosa.

Obviously, the choice of suitable carriers and other additives will depend on the exact route of administration and the nature of the particular dosage form, e.g.,

liquid dosage form (e.g., whether the composition is to be formulated into a solution, a suspension, gel or another liquid form), or solid dosage form (e.g., whether the composition is to be formulated into a pill, tablet, capsule, caplet, time release form or liquid-filled form).

- 5       Solutions, suspensions and gels, normally contain a major amount of water (preferably purified water) in addition to the antigen, lipoprotein and optional adjuvant. Minor amounts of other ingredients such as pH adjusters (e.g., a base such as NaOH), emulsifiers or dispersing agents, buffering agents, preservatives, wetting agents, jelling agents, (e.g., methylcellulose), colors and/or flavors may also be  
10       present. The compositions can be isotonic, i.e., it can have the same osmotic pressure as blood and lacrimal fluid.

- The desired isotonicity of the compositions of this invention may be accomplished using sodium chloride, or other pharmaceutically acceptable agents such as dextrose, boric acid, sodium tartrate, propylene glycol or other inorganic or  
15       organic solutes. Sodium chloride is preferred particularly for buffers containing sodium ions.

- Viscosity of the compositions may be maintained at the selected level using a pharmaceutically acceptable thickening agent. Methylcellulose is preferred because it is readily and economically available and is easy to work with. Other suitable  
20       thickening agents include, for example, xanthan gum, carboxymethyl cellulose, hydroxypropyl cellulose, carbomer, and the like. The preferred concentration of the thickener will depend upon the agent selected. The important point is to use an amount which will achieve the selected viscosity. Viscous compositions are normally prepared from solutions by the addition of such thickening agents.

- 25       As mentioned herein, a pharmaceutically acceptable preservative can be employed to increase the shelf-life of the compositions. Benzyl alcohol may be suitable, although a variety of preservatives including, for example, parabens, thimerosal, chlorobutanol, or benzalkonium chloride may also be employed. A suitable concentration of the preservative will be from 0.02% to 2% based on the total  
30       weight although there may be appreciable variation depending upon the agent selected.

      Those skilled in the art will recognize that the components of the compositions should be selected to be chemically inert with respect to the gene product and optional

adjuvant or additive. This will present no problem to those skilled in chemical and pharmaceutical principles, or problems can be readily avoided by reference to standard texts or by simple experiments (not involving undue experimentation), from this disclosure and the documents cited herein.

- 5       The inventive compositions of this invention are prepared by mixing the ingredients following generally accepted procedures. For example the selected components may be simply mixed in a blender, or other standard device to produce a concentrated mixture which may then be adjusted to the final concentration and viscosity by the addition of water or thickening agent and possibly a buffer to control  
10   pH or an additional solute to control tonicity. Generally the pH may be from about 3 to 7.5. Compositions can be administered in dosages and by techniques well known to those skilled in the medical and veterinary arts taking into consideration such factors as the age, sex, weight, and condition of the particular patient or animal, and the composition form used for administration (e.g., solid vs. liquid). Dosages for  
15   humans or other mammals can be determined without undue experimentation by the skilled artisan, from this disclosure, the documents cited herein, the Examples below.

- Suitable regimes for initial administration and booster doses or for sequential administrations also are variable, may include an initial administration followed by subsequent administrations; but nonetheless, may be ascertained by the skilled artisan,  
20   from this disclosure, the documents cited herein, and the Examples below.

- Accordingly, the invention comprehends, in further aspects, methods for preparing therapeutic compositions including a gene product or functional fragment thereof of a gene identified herein or a gene identified in an inventive method herein, as well as to methods for increasing bone density, treating, preventing or controlling  
25   osteoporosis, or otherwise alleviating a condition caused by mechanical stress or inducing bone growth, comprising administering an inventive composition, or a gene product, or functional fragment thereof of a gene identified herein or a gene identified in an inventive method herein, or a vector expressing such a gene.

- In this context and as used throughout this specification, "functional" means a  
30   protein having part or all of the primary structural conformation of the protein gene product of a gene identified herein or of a gene identified by the methods herein, and possessing the biological property of contributing to the development of bone cells in the same or an analogous fashion to the full length protein gene product, said protein



gene product being either isolated from a natural source or being the product of procaryotic or eukaryotic expression or of protein synthesis methods. The protein can have an amino acid sequence comprising an amino acid sequence of a sequence disclosed herein or of a gene product of a gene identified by a method herein or any  
5 fragment or derivative thereof by way of amino acid deletion, substitution, insertion, addition and/or replacement of the amino acid sequence. Also comprised by the term "functional" protein is the capability of said protein or part thereof to generate a specific immune response such as an antibody response; e.g., to bind to antibodies elicited by the full length protein.

- 10 Moreover, the present invention and embodiments thereof provide advances in and assist to further research and knowledge with respect to osteoporosis and conditions caused by or having as a factor mechanical stress or a lack thereof and provide an insight into development and maintenance of bone tissue. The present invention and embodiments thereof also provide advances in and assist to further  
15 clinical and epidemiological research, e.g., to allow others to further explore and extend the current potential for practical prevention and treatment. Further still, the present invention and embodiments thereof provide a deeper knowledge of factors controlling bone cell activity and regulation of bone mineral and matrix formation and remodeling contribute ultimately to the understanding of the etiology of osteoporosis  
20 or other conditions involving mechanical stress or a lack thereof. For example, the present invention provides osteoporosis or mechanical stress or lack thereof models for *in vitro* studies. This understanding will permit a more rational choice and evaluation of therapies, even as current treatments are evaluated clinically. Moreover, the present invention, for instance, via the inventive models, allows for the discovery of  
25 genes involved in processes of osteoporosis and/or one growth or formation or bone cell activity, *inter alia*. Every new gene discovered sheds more light on the complex molecular events that govern all aspects of life. The elucidation of the function of the gene and its place and role in this intricate network of pathways and structures resolves another piece in the puzzle of life. Thus, the educational and research  
30 implications are very clear. Sometimes genes may have much more benefit in this respect than in the therapeutics/diagnostics fields.

A better understanding of the present invention and of its many advantages will be had from the following examples, given by way of illustration and as a further description of the invention.

#### EXAMPLES/RESULTS

5    **Example/Result 1: ANALYSIS OF GENES AT A  
TRANSCRIPTIONAL LEVEL  
USING NUCLEAR mRNA PROBES**

#### DIFFERENTIAL TRANSLATION

10   MATERIALS/METHODS (WHICH MAY APPLY IN  
WHOLE OR PART TO SOME OR ALL EXAMPLES)

#### General Scheme

- a. Total mRNA organic extraction of all RNA from the source tissue or cell.  
(additional selection for polyA<sup>+</sup> mRNA can be included).
- b. Nuclear RNA-lysis of cells (from a tissue or a cell line) by homogenization in  
15   hypotonic buffer. Collection of nuclei by centrifugation and organic extraction of the  
RNA.
- c. Cytoplasmic RNA - Organic extraction of the RNA from the supernatant from b  
above.
- d. Polyribosomal/subpolyribosomal fractionation. Lysis of cells by homogenization  
20   hypotonic buffer, removal of nuclei and fractionation of polyribosome on linear  
sucrose gradients and organic extraction of the RNA from each fraction of the  
gradient.
- e. Secreted and membrane encoding transcripts.
  1. Isolation of RER on Percoll gradients (after homogenization of cells).
  - 25   2. Preparation of microsomes containing the RER
  3. Isolation of membrane-bound polyribosomes by successive treatment of cells  
with detergents.
- f. Nuclear proteins. Isolation of cytoskeletal associated polyribosomes by treating  
cells lyzates with different detergents.
- 30   g. Mitochondrial genes. Isolation of mitochondria on Percoll gradients.
- h. Alternative splicing. Separation of nuclei and isolation of splicosome (proteins  
and RNA complex) on linear sucrose gradients.

Preparation of cell extracts: Cells were centrifuged. The pellet was washed  
with PBS and recentrifuged. The cells were resuspended in 4x of one packed cell

- volume (PCV) with hypotonic lysis buffer (HLB: 20mM TrisHCL pH=7.4; 10mM NaCl; 3mM MgCl<sub>2</sub>). The cells were incubated five minutes on ice. 1xPCV of HLB containing 1.2% Triton X-100 and 0.2M sucrose was added. The cells were homogenized with a Dounce homogenizer (five strokes with B pestle). The cell lysate
- 5 was centrifuged at 2300g for ten minutes at 4°C. The supernatant was transferred to a new tube. HLB containing 10mg/ml heparin was added to a final concentration of 1mg/ml heparin. NaCl was added to a final concentration of 0.15M. The supernatant was frozen at
- 70°C after quick freezing in liquid N<sub>2</sub> or used immediately.
- 10 Sucrose gradient fractionation: A linear sucrose gradient from 0.5M to 1.5M sucrose in HLB was prepared. Polyallomer tubes (14X89mm) were used. 0.5 to 1.0ml of cell extract was loaded on the gradient. The cells were centrifuged at 36,000 RPM for 110 minutes at 4°C. An ISCO Density Fractionator was used to collect the fractions and record the absorbance profile.
- 15 RNA purification: SDS was added to 0.5% and Proteinase K to 0.1mg/ml and incubated at 37°C for 30 minutes. Extract with an equal volume of phenol:chloroform (1:1). The aqueous phase was extracted with one volume of chloroform and the RNA was precipitated by adding Na-Acetate to 0.3M and 2.5 volumes of ethanol and incubating at -20°C overnight. Centrifuged ten minutes, the
- 20 supernatant was aspirated and the RNA pellet was dissolved in sterile, diethylpyrocarbonate (hereinafter referred to as "DEPC") DEPC-treated water.
- Preparation of Microsomes: When possible fresh tissues and cells are used, without freezing. Tissues were powdered in liquid nitrogen with mortar and pestle and then homogenized using 4ml of buffer A/1 gr tissue (Buffer A is 250mM sucrose,
- 25 50mM TEA, 50mM KOAc pH7.5, 6mM Mg(Oac)<sub>2</sub>, 1mM EDTA, 1mM DTT, 0.5mM PMSF. PMSF was made in ethanol before making the buffer and added in drops to buffer while being stirred. This was stirred for 15 minutes and then DTT was added). Fresh organs were washed in Buffer A a few times, and then cut into pieces and homogenized. Approximately 5ml buffer A/5x10<sup>8</sup> cells were added and
- 30 homogenized. This was then homogenized on ice for 5-10 times, or as needed with the individual tissue. The mixture was transferred to 50ml tubes, then centrifuged for 10 minutes, at 4°C in a swinging bucket rotor machine. Next, the supernatant was transferred, avoiding the pellet as much as possible, to a Sorvall tube, the pellet was

washed again with 1ml buffer and centrifuge as before. The two pellets were combined, thus establishing the nuclear fraction. The combination was dissolved and treated the pellet with Tri-reagent (usually 2ml of Tri-reagent when sample is from cells) to extract the nuclear RNA. The combined 1st and 2nd supernatants were  
5 centrifuged for 10 minutes at 10000g at 4°C. Again, the supernatant was transferred to a tube and kept on ice. The pellet was washed again with 1ml buffer and centrifuged for 10 minutes at 10000g and the two pellets were combined as before, thus establishing the Mitochondrial pellet. Again, the pellet was treated with Tri-reagent (usually 1ml with cells) and the Mitochondrial RNA was extracted. Next,  
10 cold ultracentrifuge tubes were prepared containing a sucrose cushion made of: buffer A + 1.3M sucrose. The volume of the cushion was approximately 1/3 of the supernatant. The supernatant was loaded on the cushion in a 1:3 ratio of cushion to supernatant. A pair of tubes was weighed for balancing, a 20-30mg difference is allowable. The tubes were centrifuged 2.5 hours at 140,000g, 4°C with a Ti60.2 rotor  
15 (45,000 rpm). When two phases of supernatant were visible, then the red phase only was transferred (if possible), as the cytoplasmic fraction, to a sorvall tube. The clear supernatant was aspirated. When not possible to separate or phase distinction not visible, all the supernatant was taken as cytoplasmic fraction and dilute sucrose with TE (10mM Tris-HCl pH 8.0, 1mM EDTA). In the pellet were the microsomes which  
20 were visible and were clear or yellowish. For the RNA extraction, the cytoplasmic fraction was treated with 1% SDS, 0.1mg/ml proteinase K, for 30 minutes, at 37°C. After this, freezing at -80°C was possible. The RNA was extracted with a phenol:chloroform combination and precipitate with 0.3M Na-acetate, 1μl glycogen, and equal volume of isopropanol. O'N precipitation was possible and can be  
25 accomplished at 30 minutes on ice. The extract was spun at 10000g, for 20 minutes, then the RNA pellet was washed with 70% ethanol. The pellet was dried and then dissolved in H<sub>2</sub>O. The microsomes were then dissolved with 0.1M NaCl/1% SDS solution (1ml is usually sufficient for a small pellet) and extracted with a phenol:chloroform combination (no proteinase K treatment). Then the precipitation  
30 of the RNA was done in the same way as for the cytoplasmic fraction but without the requirement of adding salt.

Preparation of Nuclear and Cytoplasmic RNA: Subconfluent plates were washed with 125 mM KCl-30 mM Tris-hydrochloride (pH 7.5)-5 mM magnesium

acetate-1 mM 2-mercaptoethanol-2 mM ribonucleoside vanadyl complex (2)-0.15 mM spermine-0.05 mM spermidine at 4°C, and cells scraped from the plates were washed twice with the same buffer. Approximately  $10^8$  cells were allowed to swell for 10 minutes in 2.5 ml of swelling buffer (same as wash buffer except the KCl concentration was 10 mM) lysed with 20 strokes of a Dounce homogenizer (B pestle), overlaid on an equal volume of swelling buffer containing 25% glycerol, and centrifuged for 5 min. at 400 x g and 4°C. The upper layer of the supernatant, which contained 90% of the CAD sequences released by lysis, was designated the cytoplasmic fraction. The nuclear pellet was washed once with 2 ml of swelling buffer-25% glycerol-0.5% Triton X-100 and once with 2 ml of swelling buffer.

Nuclear RNP. Nuclei from  $10^8$  cells, prepared as described above, were suspended in 1 ml of 10 mM Tris-hydrochloride (pH 8.0)-100 mM NaCl-2 mM  $MgCl_2$ -1 mM 2-mercaptoethanol-0.15 mM spermine-0.05 mM spermidine-10 mM ribonucleoside vanadyl complex (2)-100 U of placental RNase inhibitor (Amersham Corp.) per ml and sonicated at the maximum power setting of a Konres micro-ultrasonic cell disrupter for 20 g at 4°C. Bacterial tRNA (2 mg) was added, to adsorb basic proteins (9), and the mixture was centrifuged for 1 minute (Eppendorf microcentrifuge). The supernatant was applied to a 15 to 45% sucrose gradient in mM Tris-hydrochloride-100 mM NaCl-2 mM  $MgCl_2$ -2 mM ribonucleoside vanadyl complex and centrifuged in a Beckman SW41 rotor for 90 minutes at 40,000 rpm and 4°C. RNA was recovered from gradient fractions by the addition of sodium dodecyl sulfate to 0.5%, treatment with proteinase K (200 µg/ml) for 2 hours at 37°C, extraction with phenol, and precipitation with ethanol.

Preparation of Antisense RNA: Total cellular RNA is extracted. Part of the RNA pool is immobilized on a membrane, another part converted into cDNA after ligation of oligodeoxynucleotides to the 3'-ends. The use of biotinylated, complementary oligos for cDNA synthesis allows immobilization of a "minus" strand to streptavidin-coated magnetic beads. A second set of oligos is ligated to the cDNA at the previous 5'-end of the RNA. Plus strands are eluted from the bound strands and hybridized to the membrane-bound RNA. Since the cDNA strand used has the same polarity of the RNAs, only cDNA sequences that can bind to complementary RNAs should be retained. PCR amplification and subsequent cloning of PCR-fragments is followed by sequence analysis. To test whether cloned sequences are correctly

- identified, probes are generated in sense and antisense direction. Positive clones will be structurally and functionally characterized. In order to work out this method, we started using a bacterial strain (*Escherichia coli*), containing plasmid R1 that regulates its copy number by antisense RNA. Previous work has identified both antisense
- 5 (CopA) and target RNA (CopT) of R1 intracellularly. This procedure, if feasible, will then be used to screen for antisense RNA systems in other organisms.

#### DIFFERENTIAL ANALYSIS

- Differential display: Reverse transcription: 2 $\mu$ g of RNA were annealed with 1pmol of oligo dT primer (dT)<sub>18</sub> in a volume of 6.5 $\mu$ l by heating to 70°C for five
- 10 minutes and cooling on ice. 2 $\mu$ l reaction buffer (x5), 1 $\mu$ l of 10mM dNTP mix, and 0.5 $\mu$ l of SuperScript II reverse transcriptase (GibcoBRL) was added. The reaction was carried out for one hour at 42°C. The reaction was stopped by adding 70 $\mu$ l TE (10mM Tris pH=8; 0.1mM EDTA). Oligonucleotides used for Differential display: The oligonucleotides were essentially those described in the Delta RNA
- 15 Fingerprinting kit (Clontech Labs. Inc.). There were 9 "T" oligonucleotides of the structure: 5' CATTATGCTGAGTGATATCTTTTTTTTXY 3' (SEQ ID No: ). The 10 "P" oligonucleotides were of the structure: 3' ATTAACCCTCACTAAA "TGCTGGGGA" 3' (SEQ ID No: ) where the 9 or 10 nucleotides between the parenthesis represent an arbitrary sequence and there are 10 different sequences (SEQ
- 20 ID Nos. ), one for each "P" oligo.

- Amplification reactions: each reaction is done in 20 $\mu$ l and contains 50 $\mu$ M dNTP mix, 1 $\mu$ M from each primer, 1x polymerase buffer, 1 unit expand Polymerase (Boehringer Mannheim), 2 $\mu$ Ci [ $\alpha$ -<sup>32</sup>P]dATP and 1 $\mu$ l cDNA template. Cycling
- 25 conditions were: three minutes at 95°C, then three cycles of two minutes at 94°C, five minutes at 40°C, five minutes at 68°C. This was followed by 27 cycles of one minute at 94°C, two minutes at 60°C, two minutes at 68°C. Reactions were terminated by a seven minute incubation at 68°C and addition of 20 $\mu$ l sequencing stop solution (95% formamide, 10mM NaOH, 0.025% bromophenol blue, 0.025% xylene cyanol).

- Gel analysis: 3-4 $\mu$ l were loaded onto a 5% sequencing polyacrylamide gel
- 30 and samples were electrophoresed at 2000 volts/40 milliamperes until the slow dye (xylene cyanol) was about 2 cm from the bottom. The gel was transferred to a filter paper, dried under vacuum and exposed to x-ray film.

- Recovery of differential bands: bands showing any a differential between the various pools were excised out of the dried gel and placed in a microcentrifuge tube. 50 $\mu$ l of sterile H<sub>2</sub>O were added and the tubes heated to 100°C for five minutes. 1 $\mu$ l was added to a 49 $\mu$ l PCR reaction using the same primers used for the differential display and the samples were amplified for 30 cycles of: one minute at 94°C, one minute at 60°C and one minute at 68°C. 10 $\mu$ l was analyzed on agarous gel to visualize and confirm successful amplification.

#### REPRESENTATIONAL DIFFERENCE ANALYSIS

- Reverse transcription: as above but with 2 $\mu$ g polyA+ selected mRNA.
- 10 Preparation of double stranded cDNA: cDNA from previous step was treated with alkali to remove the mRNA, precipitated and dissolved in 20 $\mu$ l H<sub>2</sub>O. 5 $\mu$ l buffer, 2 $\mu$ l 10mM dATP, H<sub>2</sub>O to 48 $\mu$ l and 2 $\mu$ l terminal deoxynucleotide transferase (TdT) were added. The reaction was incubated 2-4 hours at 37°C. 5 $\mu$ l oligo dT (1 $\mu$ g/ $\mu$ l) was added and incubated at 60°C for 5 minutes. 5 $\mu$ l 200 mM DTT, 10  $\mu$ l 10x section
- 15 buffer (100mM Mg Cl<sub>2</sub>, 900 mM Hepes, pH 6.6) 16  $\mu$ l dNTPs (1 mM), and 16 U of Klenow were added and the mixture was incubated overnight at room temperature to generate ds cDNA. 100 $\mu$ l TE was added and extracted with phenol/chloroform. The DNA was precipitated and dissolved in 50 $\mu$ l H<sub>2</sub>O.

- Generation of representations: cDNA with DpnII was digested by adding 3 $\mu$ l
- 20 DpnII reaction buffer 20 V and DpnII to 25 $\mu$ l cDNA and incubated five hours at 37°C. 50 $\mu$ l TE was added and extracted with phenol/chloroform. cDNA was precipitated and dissolved to a concentration of 10ng/ $\mu$ l.

The following oligonucleotides are used in this procedure:

- R-Bgl-12 5' GATCTGCGGTGA 3' (SEQ ID No: )
- 25 R-Bgl-24 5' AGCACTCTCCAGCCTCTCACCACA 3' (SEQ ID No: )
- J-Bgl-12 5' GATCTGTTTCATG 3' (SEQ ID No: )
- J-Bgl-24 5' ACCGACGTCGACTATCCATGAACA 3' (SEQ ID No: )
- N-Bgl-12 5' GATCTTCCTCG 3' (SEQ ID No: )
- N-Bgl-24 5' AGGCAACTGTGCTATCCGAGGGAA 3' (SEQ IDNo: )
- 30 R-Bgl-12 and R-Bgl-24 oligos were ligated to Tester and Driver: 1.2 $\mu$ g DpnII digested cDNA. 4 $\mu$ l from each oligo and 5 $\mu$ l ligation buffer X10 and annealed at 60°C for ten minutes. 2 $\mu$ l ligase was added and incubated overnight at 16°C. The

- ligation mixture was diluted by adding 140 $\mu$ l TE. Amplification was carried out in a volume of 200 $\mu$ l using R-Bgl-24 primer and 2 $\mu$ l ligation product and repeated in twenty tubes for each sample. Before adding Taq DNA polymerase, the tubes were heated to 72°C for three minutes. PCR conditions were as follows: five minutes at 72°C, twenty cycles of one minute at 95°C and three minutes at 72°C, followed by ten minutes at 72°C.

Every four reactions were combined, extracted with phenol/chloroform and precipitated. Amplified DNA was dissolved to a concentration of 0.5 $\mu$ g/ $\mu$ l and all samples were pooled.

- 10        Subtraction: Tester DNA (20 $\mu$ g) was digested with DpnII as above and separated on a 1.2% agarous gel. The DNA was extracted from the gel and 2 $\mu$ g was ligated to J-Bgl-12 and J-Bgl124 oligos as described above for the R-oligos. The ligated Tester DNA was diluted to 10ng/ $\mu$ l with TE. Driver DNA was digested with DpnII and repurified to a final concentration of 0.5 $\mu$ g/ $\mu$ l. Mix 40 $\mu$ g of Driver DNA  
15        with 0.4 $\mu$ g of Tester DNA. Extraction was carried out with phenol/chloroform and precipitated using two washes with 70% ethanol, resuspended DNA in 4 $\mu$ l of 30mM EPPS pH=8.0, 3mM EDTA and overlayed with 35 $\mu$ l mineral oil. Denatured at 98°C for five minutes, cool to 67°C and 1 $\mu$ l of 5M NaCl was added to the DNA. Incubated at 67°C for twenty hours. Diluted DNA by adding 400 $\mu$ l TE.

- 20        Amplification: Amplification of subtracted DNA in a final volume of 200 $\mu$ l as follows: Buffer, nucleotides and 20 $\mu$ l of the diluted DNA were added, heated to 72°C, and Taq DNA polymerase was added. Incubated at 72°C for five minutes and added J-Bgl-24 oligo. Ten cycles of one minute at 95°C, three minutes at 70°C were performed. Incubated ten minutes at 72°C. The amplification was repeated in four  
25        separate tubes. The amplified DNA was extracted with phenol/chloroform, precipitated and all four tubes were combined in 40 $\mu$ l 0.2XTE. Digested with Mung Bean Nuclease as follows: To 20 $\mu$ l DNA 4 $\mu$ l buffer, 14 $\mu$ l H<sub>2</sub>O and 2 $\mu$ l Mung Bean Nuclease (10 units/ $\mu$ l) was added. Incubated at 30°C for thirty-five minutes + First Differential Product (DPI).

- 30        Repeat subtraction hybridization and PCR amplification at driver: differential ratio of 1:400 (DPII) and 1:40,000 (DPIII) using N-Bgl oligonucleotides and J-Bgl



oligonucleotides, respectively. Differential products were cloned into a Bluescript vector at the BAM HI site for analysis of the individual clones.

The experimental cells were grown alternatively under normal conditions, for 4 hours under hypoxia (<1% oxygen) and for 16 hours under hypoxia. The cells were harvested and RNA was extracted either from nuclei that were prepared from the cells (nuclear RNA) or from extracts of unfractionated cells (total cellular RNA).

Figure 2 demonstrates how the probes prepared from the nuclear RNA (STP) give a higher differential expression than the total cellular RNA probe (Tot). The control genes encoding VEGF (vascular endothelial growth factor), Glut1 (glucose transporter 1) and glycogen synthase are known to be induced by the hypoxia stress. The level of induction observed in the nuclear probe is much higher than that seen in the total probe and much closer to the actual know level of induction. The three new genes RTP 241, RTP 262 and RTP 779 show marked induction by hypoxia. Again, the induction level seen with the nuclear probe is much higher, up to five-fold higher, as seen for RTP779. When the induction of these genes was analyzed by the Northern blot method, it was found that the nuclear probe was once again much closer to the actual situation, while the total probe gives a marked underestimation.

The genes RTPi-66 and RTP21-72 demonstrate the ability of the nuclear probe to detect differentially expressed genes that do not appear differentially with the total probe.

The genes for Nucleolin and Thrombospondin show that also for down-regulated mRNAs the nuclear probe is much more sensitive and gives much high levels of differential expression values.

Lastly, the genes for ribosomal protein L17 and cytoplasmic gamma-actin are known as genes that do not respond to hypoxia stress. The nuclear probe and the total probe both show that no induction occurs.

#### **Example/Result 2: DIFFERENTIAL EXPRESSION PROFILING**

Chip: The microarray (Chip) used was prepared as follows. Subtraction experiments

were carried out on rat osteoblasts (Calvaria) using CLONTECH SSH kit

(K 1804- 1). Cells were subjected to 20 minutes of mechanical stress and compared to "normal" cells not subjected to mechanical force. 767

induced sequences and 606 reduced sequences were selected and printed on a chip.

Probe: Total RNA

Cells: Primary Calvaria cultures derived from 17-19 days old rat embryos.

5 List of Analyses:

CHIP no	Experiment	Analysis
109	3 -Ca +/-mechanical force	Compare the system with and without application of mechanical force in absence of Ca in culture medium. Find genes differentially expressed under the influence of mechanical force.
110	2 +Ca +/-mechanical force	Compare the system with and without application of mechanical force in presence of Ca in culture medium. Find genes differentially expressed under the influence of mechanical force.
111	1 +Indomethacin +/- mechanical force	Compare the system with and without application of mechanical force. Find genes differentially expressed under the influence of mechanical force (with prostaglandin synthesis inhibited by indomethacin).
107	4 +/- PEG2	Compare the system with and without PGE2 treatment (to mimic mechanical force)
116	5 +/- PEG2	Compare the system with and without PGE2 treatment (to mimic mechanical force)

**Calvaria treated with indomethacin and mechanical force**

Primary cell cultures derived from 17-19 days old rat embryos. The cultures were prepared by trypsin - EDTA digestion of Calvaria including the periosteum. The  
 10 cell cultures were grown in MEM medium with 10% FCS for 5-6 days to reach confluency.

At this time 10 microliter which contains 20 $\mu$ g of indomethacin were added to culture dishes which had 4ml of medium. 20 minutes later the dishes were activated mechanically. The mechanical activation is carried out e.g. by expanding an  
 15 orthodontic expansion screw which is attached to two pieces of solid acrylic resin

glued to the outer surface of the cell culture dish. The expansion deforms the dish irreversibly. Same cultures which were not treated with indomethacin and activated mechanically were as positive control.

- 5 The rationale is because mechanical activation stimulates *de novo* synthesis of prostaglandins. Indomethacin inhibits synthesis of prostaglandins.

**Calvaria, grown in the presence of Ca activated by mechanical force**

- 10 Calvaria grown in the presence of Ca were activated by mechanical activating devices at confluency. The cultures were prepared as described above. The cells were grown in MEM medium which normally includes 1 mM of calcium: - from the seeding of the cells until confluency and mechanical activation.

**Calvaria grown in absence of Ca treated with mechanical force**

- 15 The cultures were grown in MEM medium which was calcium free. The calcium in this medium was 0.25 mM because it consisted of 10% FCS (serum contains 2.5 mM of calcium). After 3-4 days the medium was replaced by regular MEM which included normal calcium concentration.

- 20 The rationale of this experiment is mainly because transduction mechanism of mechanical activation like prostaglandin synthesis and action is calcium dependent. The cultures in low calcium suppress the proliferation of fibroblasts and allow growth and differentiation of osteoblasts in culture. Therefore, the strategy is to start with low calcium medium and after 3-4 days to booster growth by switching to normal calcium medium (1 mM Ca).

**Calvaria treated with prostaglandin PGE2**

- 25 Primary Calvaria cells were treated with PGE2 in both experiments (both are similar primary cultures prepared at different date, and treated identically).

PGE2 treatment was performed on cultures which reached confluency by adding (treating with) 10 microliter of PGE2 which consists of total of 500ng of PGE2. After 30 minutes the cells were scraped and stored in -70°C.

The rationale: The PGE2 treatment is supposed to mimic the mechanical activation effect.

- 30 The results from this Example are shown in the Table and sequences of Figure 2A. Novel sequences CMF608, CMF405 and CMF274 were identified, *inter alia*, as discussed below (see also Figs 2A-14).

**AN OSTEOPOROSIS (or mechanical stress) MODEL (calvaria cell cultures).****Differentially regulated/differentially expressed genes post-mechanical stimulation****Extracellular matrix, transmembranal and secreted proteins:**

- 5    tenascin
- collagen XII
- thrombospondin 1
- ADAMTS-1.
- C3 complement component
- 10   alpha-2-macroglobulin receptor
- fibronectin
- connective tissue growth factor
- endothelin converting enzyme
- alpha-2u microglobulin-related protein
- 15   RB13-6

**Genes connected to regulation of apoptosis**

SARP1  
cytochrome oxidase subunit 1  
glutamyl-cystein synthetase

- 20   **Genes connected to intracellular fatty acid methabolism**
- 3-hydroxy-3-methylglutaryl coenzyme A reductase
- yeast ERG3 homologue
- and yeast ERG25 homologue
- stearoyl-CoA desaturase

- 25   **Genes connected to cytoskeleton regulation**

AHNAK  
filamin  
syntrophin 1

**Genes connected to regulation of water channels**

- 30   aquaporin1

**Novel genes or known anonymous genes without function**

- highly charged amino acid sequence
- DEST274 (CMF274; see Figs 2A-14)
- DEST405 (CMF405; see Figs 2A-14)
- 35   DEST608 (CMF 608; see Figs 2A-14)

**General overview of identified genes.**

- Tenascin is an extracellular matrix glycoprotein whose expression is up-regulated in normal bone development during condensation. It is also involved in genesis and function of articular chondrocytes. Tenascin is secreted by osteoblasts,
- 40   but is absent from mineralized bone matrix. Expression of alkaline phosphatase

activity and **collagen XII** (markers of osteoblast differentiation) are tenascin-dependent (down-regulated by anti-tenascin RNA). Expression of tenascin is markedly increased in response to mechanical stress, its promoter (in chicken) was shown to contain a cis-acting "strain-responsive" element.

- 5           Another protein whose expression is known to be modulated by mechanical stress (fluid shear or stretched stress of i.e. mesothelial cells) is endothelin (**endothelin converting enzyme** that generates active endothelin molecules from inactive intermediates is upregulated in the present screen). In bone, endothelin stimulates the osteoblastic IL-1-induced production of IL-6 - mediator of osteoclastic
- 10   differentiation, function and probably survival. Receptors to endothelin were demonstrated in osteoblastic cells by ligand binding (autocrine loop). Major endothelin signal transduction pathways in bone cells is stimulation of phospholipid turnover, by activation of phospholipases A, C, and D, stimulation of Ca flux from intra- and extracellular stores and activation of tyrosine kinases. Endothelins also
- 15   modulated calcium signaling elicited by other agents (i.e. potentiation of PH-stimulated Ca transient) in osteoblastic cells. Phenotypic responses to endothelin include stimulation of osteocalcin and osteopontin messages (see herein), inhibition of osteoclast motility and stimulation of prostaglandin-dependent resorption.

- One protein exhibiting channel characteristics was found upregulated. It is
- 20   **aquaporin1** - a water channel protein expressed in many fluid secreting and absorbing tissues such as kidney, brain, heart, eye, inner ear. Its promoter contains glucocorticoid responsive elements and can be activated in response to dexamethasone treatment. Induction of aquaporin-1 expression was detected by subtracted cloning of genes upregulated following cardiopulmonary bypass and
- 25   reperfusion. However, its induction is delayed compared to inflammatory mediators (i.e. ICAM-1, E-selectin, IL-8). The only bone link can be traced in localization of aquaporin molecules in the inner ear, but this localization can be easily explained by the critical dependence of inner ear function on fluid homeostasis. In ear, the protein was found in close association with bone - in most of the cells lining the bony
- 30   labyrinth, and in other non-bony locations.

**AHNAK** (other names: neuroblast differentiation factor, desmoykin) - a 700 kD protein that was originally identified as differentially repressed (lost) in neuroblastoma cells. Its body is constituted of 128 amino acid repeats. The protein

was initially identified as a nuclear one. However, when it was rediscovered under the name "desmoykin", its subcellular localization was reported as membranous (at the sites of desmosomes). AHNAK-like repeats were found in another protein VAP-1 (vesicle associated protein) - a novel high molecular weight protein found in sea urchin eggs. It is located at peripheral membrane in association with microsomal membrane fraction. Within AHNAK-like repeats of VAP-1 RNA-binding sequences - of RNP1 and of RNP2 types (the same is true for AHNAK). Therefore, it is tempting to speculate, that the general increase in expression of secreted proteins observed in bone tissue in response to mechanical stress might dictate the need in the upregulation of RNA-binding protein localized to a microsomal fraction.

**Filamin** (non-muscle type), ABP-280, plays a critical role in stabilizing the membrane-cytoskeletal interactions. It is a dimeric actin crosslinking protein that provides the major mode for attaching the cortical F-actin network to membrane glycoproteins. One filamin molecule is able to crosslink up to 1,000 actin molecules. This ability makes filamin the most potent actin crosslinking agent known today.

**Syntrophin 1** is a member of multigene family of intracellular extrinsic membrane proteins found in complex with dystrophin. This particular syntrophin was demonstrated also in complex with nitric oxide (NO) synthase (in muscle tissue). The interaction is likely to be mediated by PDZ domains found in both proteins, but formation of this complex is probably dystrophin-dependent. NO is known to be implicated in the metabolism of bone, especially as a mediator of cytokine effects on remodeling of bone tissue in response to diverse stimuli such as pro-inflammatory cytokines, mechanical stress and sex hormones. Both estrogen and mechanical stress increase NO production by activating constitutive nitric oxide synthase. High concentrations of NO inhibit bone resorption by inhibiting osteoclast formation and by inhibiting the resorptive function of mature osteoclasts, whereas lower NO concentrations potentiate bone resorption and may be essential for normal osteoclast function. On the other hand, growth and differentiation of osteoblasts are also inhibited by high NO concentration.

**Thrombospondin 1** - a 450 kD adhesive glycoprotein involved in cellular attachment, spreading, proliferation, and migration. It was originally isolated from platelets and endothelial cells, but it is also localized in osteoid of undermineralized fetal subperiosteum and in mineralized bone matrix of neonatal/young (growing)

bone. TSP-1 can specifically interact with osteonectin - a 30 kD protein of bones and platelets. This complex formation is Ca-dependent. In osteogenesis imperfecta, levels of osteonectin are reduced, while production of thrombospondin is increased. Expression of thrombospondin is a marker of osteoblast differentiation (together with alkaline phosphatase and alpha-1-collagen). Dexamethasone treatment decreases the levels of thrombospondin expression in cultured osteoblastic cells (glucocorticoids induce osteoporosis). 17-beta estradiol, on the contrary, induces thrombospondin expression. Thrombospondin-1 gene expression is modulated during pericytes differentiation in vitro (pericytes are cells that are embedded within basement membrane of microvessels, believed to participate in angiogenesis, but are able to differentiate into osteogenic phenotype). It is markedly increased during nodule formation and then decreased when mineralization of the nodules has taken place. TSP-1 is excluded from the inner mass of such mineralized nodules.

Several non-thrombospondin genes were found to contain **thrombospondin motifs** (cell-binding domain of thrombospondin). One of them also belongs to a metalloproteinase-disintegrin family (identified as an up-regulated gene in the present screen) - **ADAMTS-1**. It was initially cloned as a gene that is selectively expressed in cachectic (in vivo) colon 26 adenocarcinoma subline. It is a putative secreted protein without transmembranal domain. ADAMTS-1 contains six protein modules: pro-, metalloproteinase, disintegrin-like. TSP type 1 motif, spacer, C-terminal TSP motifs.

Another TSP-motifs containing protein is properdin - a plasma glycoprotein which stabilizes the C3nBb enzyme complex of the alternative pathway of the complement system through TSP motifs binding. Interestingly these motifs are also found in terminal complement components C6 - C9.

**C3 complement component** is produced by osteoblastic and marrow-derived stromal cells in response to vitamin D and regulates differentiation of mononuclear phagocytes into osteoclasts. This effect is bone-specific, since C3 serum, unlike bone, concentrations were unaffected in vitD-deficient mice. In normal mice the C3 protein is located mainly in periosteal regions of calvaria and on the surfaces of bone trabeculae in tibial metaphyses. It is suggested that C3 deposition on mineralized bone surfaces mediates recruitment of mononuclear osteoclasts (unlike multinuclear, express C3 receptor) to this site. In biological fluids, activated C3 in complex with

**alpha-2-macroglobulin** (whose receptor was found to be upregulated in this screen - this **receptor** is known to be expressed by bone marrow macrophages, so, probably, osteoclast precursors can be alpha-2M-receptor positive as well) binds IL-1. It is worth noting, that IL-1 is considered as one of the stimulators of osteoclastogenesis and treatment of ovariectomized mice with its inhibitor significantly decreases the bone loss. Increased osteoclast development after estrogen loss is also mediated by IL-6. Both cytokines expression is upregulated in vivo and in vitro following estrogen deprivation.

It seems now proven that estrogen induces apoptosis of bone-resorbing osteoclasts being applied directly. On the other hand, estrogen induces TGF-beta 1 production by osteoblasts, and anti-TGF-beta antibodies, in turn, can inhibit the estrogen-induced apoptosis of osteoclasts. In this light, finding of **SARP1** upregulation is of special interest. SARPs are a family of secreted apoptosis-related proteins. **SARP1** was initially identified as a component of conditioned medium collected from quiescent cells, responsible for apoptosis resistance. **SARP2**, on the contrary, induces apoptosis sensitization. Structurally SARPs possess a cystein-rich domain (CRD), homologous to CRD of frizzled proteins, but lack the transmembranal domain.

Upregulation of **cytochrom oxidase subunit 1** may be a consequence of mechanical stress or oxidative stress/apoptosis possibly mediated in the system by, for example, increased NO levels.

**DEST (ACC#AA177798)**, after the contige construction turned out to belong to a cDNA coding for **glutamyl-cystein synthetase** - a rate limiting enzyme in glutathione (GSH) synthesis. Its upregulation may be related to the stressed conditions like in the previous case. On the other hand there is one clinical work that correlates GSH reduction (low activity of antioxidant systems) in patients with hypomineralized state of bones.

TGF-beta1 is known as a principal inducer of **connective tissue growth factor** (CTGF, cef10, flsp12, cyr61, beta1G-M1, beta IG-M2, nov-protoncogene) expression. The latter contains four distinct structural modules, each of them being homologous to distinct domains in other extracellular proteins such as Von Willebrand factor, slit, **trombospondins**, fibrillar collagenes, IGF-binding proteins and mucins. CTGF expression is induced not only by TGF-beta1, but also by **BMP2**



(bone morphogenic factor 2), and during wound repair. In embryogenesis, its expression is found in developing cartilaginous elements, including limbs, ribs, prevertebrae, chondrocranium and craniofacial elements (Meckel's cartilage). Thus, CTGF transcription correlates with differentiation of chondrocytes of both mesodermal and ectodermal origin. In culture, CTGF is expressed in chondrocytes but not in osteoblasts. Possible role in endochondral ossification is suspected because of responsiveness to BMP2. In fibroblasts, CTGF expression causes upregulation of alpha-1-collagen, alpha-5-integrin and fibronectin.

- Several enzymes known to participate in steroid synthesis were found transcriptionally elevated in the present system in response to mechanical stress. They include 3-hydroxy-3-methylglutaryl coenzyme A reductase (the first rate limiting enzyme in the chain of cholesterol synthesis from 3 acetyl-CoA molecules), yeast ERG3 homolog - sterol-C5-desaturase and yeast ERG25 homolog - methyl-sterol-oxidase (both may play a role in formation of cholesterol from lanosterol). It is worth noting that cholesterol is the basis for estrogen and vitamin D3 synthesis. One additional enzyme belonging to fatty acid metabolic pathways that was found upregulated is stearoyl-CoA desaturase, that converts the saturated substrate into the D9-deaturated oleoyl-CoA. Both compounds participate in the synthesis of phospholipids building the cell membrane. Interestingly, estrogens and androsterons are known enhancers of the desaturation reaction.

- Alpha-2u microglobulin-related protein** (neutrophil gelatinase-associated lipocalin precursor - NGAL) belongs to a lipocalin superfamily embracing small extracellular proteins that can bind small hydrophobic molecules (i.e. retinols) and serve ligands to specific extracellular receptors. Many of them were implicated in regulation of cell homeostasis. NGAL was identified as a protein secreted from specific neutrophils' granules upon cell activation and it is identical to a 24p3 protein upregulated in SV-40 induced mitotic reaction. Interestingly, NGAL expression is increased in neu- but not in ras-induced experimental mammary tumors. NGAL can be upregulated by dexamethasone through a responsive promoter element in vitro. In vivo, induction of NGAL in epithelial cells was observed in inflammatory and neoplastic colorectal diseases, but not in normal colon. Among other lipocalins, NGAL is mostly similar to lipocalin-type prostaglandin-D-synthase, responsible for synthesis of prostaglandin D2 from prostaglandin H2. However, NGAL is not

supposed to have the enzymatic activity because of the absence of a specific Cys residue (position 65) which is crucial for prostaglandin-D-synthase function. Structural similarity between two proteins most probably stems from clustered localization of both genes at the same chromosomal locus. Nothing is known about

5 NGAL function in bones.

In addition, two known (RB13-6 antigen and highly charged amino acid sequence ACC# X59131) but without any attributed function and three novel proteins designated 274, 405 and 608 were found to be upregulated in the present model.

- 10 **RB13-6** is a cell surface 130 kD glycoprotein selectively recognized by monoclonal antibody with the same name. RB13-6 is as a surface antigen of a subset of glial cells highly susceptible to malignant conversion by treatment with a certain carcinogen. This protein is related to the human and murine plasma cell membrane protein PC-1, a nucleotide pyrophosphatase / alkaline phosphodiesterase, and possesses a 5'-nucleotidase activity. However, unlike PC-1, RB13-6 contains an
- 15 RGD-sequence. The latter is a signature of integrin-interacting proteins.

- So called **highly charged amino acid sequence (ACC#X59131)** is a putative protein encoded by anonymous open reading frame of 315 amino acids. It has no significant homology to any protein in the database. Charged amino acids are found in clusters with either Ser and Thr or Ser and Pro residues. Two prominent alpha-helices
- 20 - one basic and one acidic - are positioned near the C-terminus.

- 274 (novel gene):** In rat calvaria primary cell cultures, expression of this gene was found upregulated approximately 3-fold by mechanical strain. This was detected both by microarray analysis and by Northern hybridization. In rat calvaria this gene is expressed as a single RNA species of approximately 9 Kb. However, expansion of
- 25 Northern analysis to RNA samples from other rat tissue sources we have found that 274 may probably be alternatively spliced in a tissue specific manner. Alternatively, there is a family of genes closely related to 274 genes that are differently expressed in different tissues. Transcripts of varying length (in general, 9 Kb or a slightly shorter) were found in rat small intestine, skeletal muscle, lung, kidney, eye, brain, colon and
- 30 testis. The highest expression levels were found in testis, eye and kidney. Complex expression pattern was discovered in bone: two strong transcripts of more than 9 Kb and 4 Kb and three faint transcripts of 1.8 Kb, 0.5 Kb and 0.3 Kb. Interestingly, when human derived lymphoid cell line NB4 was hybridized to the same probe, these three

faint transcripts appeared strong, while bone-specific two strong transcripts were not evident. This, probably, indicates that the origin of the three short transcripts is in some lymphoid precursors present in small amount in the bone marrow.

cDNA library was prepared from RNA extracted from calvaria cells after mechanical stimulation. 5291 bp 274-specific RACE product was synthesized and sequenced (see Figs 2A-14). Comparison to public databases revealed that 274 is a rat homologue (98% identity on the level of amino acids) of anonymous human cDNA KIAA0462. This sequence is 7150 bp long and contains open reading frame of 6900 bp capable of coding for 2300 amino acid protein. The frame is still open from the 5', indicating the lack of the N-terminal sequences. The open reading frame extends to the 5' direction for additional 900 bp compared to the human KIAA0462 sequence, but still does not reach the beginning of the protein. On the basis of human sequence information the inventors were able to synthesize a 6.8 Kb long human specific RACE contig. The putative KIAA0462 protein has no direct analogs in the database. It distantly (26%) resembles the *C.elegans* hypothetical protein (AF003140) that was, in turn, defined as having weak similarity to the drosophila hyperplastic disc protein. No known protein functional domains were identified either. A stretch of 24 hydrophobic amino acids between positions 165 and 188 of the KIAA0462 putative protein hints on its potential transmembranal location.

In situ hybridization analysis (discussed in more detail in further Examples, *infra*): in normal rat bones and bones obtained from ovariectomized (osteoporotic) rats gave preliminary results indicating that gene 274 is expressed in long bones of normal rats in lining cells covering the inner surface of compact bone and in bone marrow. In normal trabecular bone specific signal was detected in bone marrow. Osteoporotic bones were 274 - negative or displayed extremely reduced signal.

CMF274 appears to encode a huge protein. This gene seems to be bone specific. In the bone expression is in many compartments. Expressed in osteoblasts but not in lining cells. This gene seems different from the other two in that it is expressed in more mature cells.

With respect to CMF 274, mention is made of Xu et al, "Retinal Targets for Calmodulin Include Proteins Implicated in Synaptic Transmission," J Biol Chem 273(47):31297-307 (1998). Xu relates to what may be a homolog for CMF274 in drosophila, named "calossin". Xu et al. may provide homology to mouse ESTs and

that the protein is highly conserved during evolution. The mouse ESTs may be from the mouse homolog of CMF274. However, in Xu et al., the mouse gene was not characterized, and Xu et al. do not provide any relation to bones.

- CMF274 contains some interesting domains including calmodulin binding  
5 domain and two zinc finger domains. The first implies capability to bind the important Calcium sensor calmodulin, and the second implies DNA binding capabilities. This combination support the proposition that CMF274 is involved in a central aspect of bone biology: sensing amount of calcium and translating it into nuclear signals that change the expression of downstream genes. It is important to note that one of the  
10 zinc-finger binding domains (CRD1) "is most similar to the zinc-finger family defined by Requiem, a protein required for apoptosis" (Xu et al).

- 405 (novel gene):** In rat calvaria primary cell cultures, expression of this gene was found to be downregulated in response to mechanical stimulation as detected both by microarray analysis and by Northern blots. A single 9 Kb transcript was detected  
15 in this tissue. However, being hybridized to rat tissue blot, the same probe had detected a major 5 Kb transcript ubiquitously expressed in bone, brain, colon, small intestine, testis, ovary, uterus, heart, kidney, liver, stomach, thymus, spleen, bladder, adipose tissue and mammary gland.

- Partial 405 rat cDNA clone was isolated from rat calvaria cDNA library by  
20 RACE technique and sequenced. It contains 3684 bp, 3000 of them constituting an open reading frame closed from the 3' end (see Figs 2A-14). Comparison to public databases revealed that gene 405 is a rat homologue of human anonymous cDNA sequences KIAA0183 and AF055017. Interestingly, the latter cDNA lacks a 294 bp (98 amino acids) fragment corresponding to positions 2384 - 2483 of KIAA0183  
25 putative amino acid sequence. Rat 405 homologue contains this region.

- Thus, there is an indication that gene 405 is subjected to alternative splicing. KIAA0183 cDNA clone was obtained from Genbank and used as a probe for hybridization with human tissue RNA blot. As in the case of rat tissues, a 5 Kb transcript was detected. However, in humans its expression was not as uniformly  
30 distributed among different tissues as in rats. The highest levels were detected in early embryo, and in testis, placenta, ovary, tongue, intestine in adults. The only RNA of human origin where a faint 9 Kb transcript (together with the major 5 Kb one) could be seen was RNA from K562 early myeloid precursor cell line.

KIAA0183 cDNA clone represents a 1062 amino acid open reading frame, lacking the N-terminus. The available sequence has no transmembrane domain. On the other hand, four structural subdomains can be easily identified: N-terminal, highly charged alpha-helical region, Ser-Pro rich spacer domain, C-terminal highly charged alpha-helical region, and a tail region, rich in Ser, Pro, Gly, and Arg - amino acids, known to be clustered in this composition in the RNA-binding protein regions. A middle spacer region contains an RGD motif known to serve as a receptor to integrins. No significant homology to any known protein was found. Among those ones that displayed distant fragment homology, are FUS/TLS RNA binding protein (nuclear export) and various types of collagens.

Thus, as CMF405 the full sequence of the human cDNA is also presented. The pattern of expression in bone suggests involvement in osteoblast and chondrocyte differentiation. The presence of an RGD motif in this protein suggests involvement in response to integrins. Expression in few other tissues may suggest a broader function.

**608 (novel gene):** In rat calvaria primary cell cultures, expression of this gene was found upregulated approximately 3-fold by mechanical strain. This was detected both by microarray analysis and by Northern hybridization. In rat calvaria this gene is expressed as a single RNA species of approximately 9 Kb. Hybridization signal was not detected in any other rat RNA from different tissue sources, including testis, colon, intestine, kidney, stomach, thymus, lung, uterus, heart, brain, liver, eye, and lymphnode.

Partial 608 rat cDNA clone was isolated from rat calvaria cDNA library by RACE technique and sequenced. RACE contig is 4007 bp long and contains a 3356 bp open reading frame closed from the 3'. Comparison to public databases revealed no sequence homologues. There are several human EST clones, similar 608 cDNA. The primary structure of the putative protein enables to attribute it to Ig superfamily.

By *in situ* hybridization (discussed in more details in Examples *infra*), expression of gene 608 was found in bone marrow from normal trabecular bones, but not of osteoporotic ones.

As to CMF608, the inventors found that it encodes a big protein that is most probably a part of the extra-cellular matrix. The gene may be actively involved in supporting osteoblast differentiation. Another option is that it marks regions were

remodeling takes place. Such an hypothesis is also compatible with a role in directing osteoclast action and thus it may be a target for inhibition by small molecules.

In normal bone formation, activation of osteoblasts leads to secretion of various factors that attract osteoclast precursors or mature osteoclasts to the sites of bone formation to initiate the process of bone resorption. In normal bone formation both functions are balanced. Imbalance to any side causes either osteopetrosis (osteoblast function overwhelms) or osteoporosis (osteoclast function overwhelms).

Among known osteoblast activators - mechanical force stimulation - is actually applied in the present model. As proof of principle, increased expression of several genes known to respond to mechanical stress by transcriptional upregulation were found. They include tenascin, endothelin and possibly thrombospondin. Upregulation of water channel encoding message is likely related to this mechanism too.

Among genes whose expression is found upregulated are those known to be expressed by activated osteoblasts (i.e. complement C3) or those whose upregulation may be logically connected:

1) Mechanical stress activates constitutive NO-synthase (NO plays an important role in bone building). In muscle, this enzyme is found in complex with syntrophin 1. The latter is found upregulated in our screen. Therefore, it can participate in NO-synthase activation in bone, too.

2) Some proteins secreted by osteoblasts participate osteoclast attraction through the RGD-mediated binding to integrin receptor highly expressed by osteoclasts, e.g., osteopontin. While this specific protein was not identified in the present screen, this is of no moment as other RGD-containing proteins were identified in the present screen, including the DEST405 and RB13-6.

3) Additional protein complex that attracts osteoclasts is the complex of C3 complement. Such complexes are known to be stabilized in plasma by a thrombospondin motifs containing glycoprotein - properdin. Another thrombospondin motifs-containing protein without any known function - ADAMTS-1 is found upregulated. It can participate in stabilization of C3 complexes in bone.

Several lines of evidence indicate that mechanical stress causes generation of both apoptotic (NO, cytochrom oxidase subunit 1) and antiapoptotic signalling

(SARP1, glutamyl-cystein synthetase). That may be important for keeping the balance between osteoblast and osteoclast proliferation, differentiation and death.

Interestingly, the inventors found that in response to mechanical force, several enzymes that regulate the chain of chemical reactions potentially leading steroid

5 synthesis are upregulated. In bone, these steroids may have the estrogen-like function that is reflected by at least two observations: first - upregulation of stearyl-CoA-desaturase (elevation of this enzyme activity is known as estrogen-dependent); and, second - upregulation of connective tissue growth factor that is mainly induced in cells by TGF-beta -1, that, in turn, is known to be induced in osteoblasts by estrogen.  
10 Such a link may explain a common anti-osteoporotic action of estrogens and mechanical force. It is worth noting, that estrogen induces osteoclasts apoptosis both directly (applied to cultured osteoclasts) and in TGF-beta-1-dependent manner.

Genes that are implicated in osteoporosis, osteoporosis prevention, treatment, or control, or in study or investigation to advance knowledge of osteoporosis, its

15 prevention, treatment or control and/or in bone growth/formation study or investigation and/or for addressing maladies, conditions, symptoms, and the like, associated with bone growth/formation, include three novel genes DEST's 274, 405, 608, as well as RG13-6 (as RGD-containing protein), metalloproteinase ADAMTS-1, and proteins of SARP family (secreted apoptosis related proteins) including as  
20 potential modifiers of programmed cell death in bone formation.

Furthermore, as mentioned herein, this Example can also be performed without imparting mechanical stress to the cells; for instance, in reduced or zero gravity conditions, to develop a model with the lack of mechanical stress; and, the invention comprehends such a model and genes thereby identified.

25 **Example/Result 3: CMF608 GENE EXPRESSION BY *IN SITU* HYBRIDIZATION**

Pattern of expression of CMF608 gene was studied by in situ hybridization on sections of bones from ovariectomized and sham-operated rats. Female Wistar rats weighting 300-350 g were subjected to ovariectomy under general anesthesia. Control  
30 rats were operated in the same way but ovaries were not excised - sham operation.

Three weeks after operation rats were sacrificed and tibia were excised together with the knee joint. Bones were fixed for three days in 4% paraformaldehyde and then decalcified for four days in solution containing 5% formic acid and 10%

formalin. Decalcified bones were postfixed in 10% formalin for three days and embedded into paraffin.

To study the pattern of expression of CMF608 gene in developing bone the model of ectopic bone formation was employed. Rat bone marrow cells were seeded into cylinders of demineralized bone matrix prepared from rat tibiae. Cylinders were implanted subcutaneously into adult rats. After three weeks rats were sacrificed and implants were decalcified and embedded into paraffin as described above for tibial bones.

The 6  $\mu$ m sections were prepared and subjected to *in situ* hybridization procedure. After hybridization sections were dipped into nuclear track emulsion and exposed for three weeks at 4°C. Autoradiographs were developed, stained with hematoxylin-eosin and studied under microscope using brightfield and darkfield illumination.

For the further assessment of cell and tissue specificity of CMF608 gene expression *in situ* hybridization study was performed on sections of multitissue block containing multiple samples of adult rat tissues. Developmental pattern of CMF608 expression was studied on sagittal sections of mouse embryos of 12.5, 14.5 and 16.5 days postconception (dpc) stages.

Microscopic study of hybridized sections of long bones revealed a peculiar pattern of CMF608 probe hybridization. The hybridization signal can be seen mainly in fibroblast-like cells found in several locations throughout the sections. Prominent accumulations of these cells can be seen in the area of periosteal modeling in metaphysis, and also in regions of active remodelling of compact bone in diaphysis: at the boundary between bone marrow and endosteal osteoblasts and in periosteum, also in close contact with osteoblasts. Perivascular connective tissue filling Volkmann's canals in compact bone in diaphysis and epiphysis also contains expressing cells. No hybridization is found within cancellous bone and in bone marrow. This pattern of hybridization suggests that cells showing expression of CMF608 are associated with areas of remodelling of preexisting bone and are not involved in primary endochondral ossification.

At the level of growth plate expressing cells can be seen in perichondral fibrous ring of LaCroix. Some investigators regard this fibrous tissue as the aggregation of residual mesenchymal cells able to differentiate into both osteoblasts



and chondrocytes. In this respect it is noteworthy that single cells expressing CMF608 can be seen in epiphyseal cartilage. These expressing cells are rounded cells within the lateral segment of epiphysis (sometimes in close vicinity to the ring of LaCroix) and flattened cells covering the articulate surface. Most of cells in articulate cartilage and all chondrocytes of growth plate do not show expression of CMF608. Ovariectomy did not result in change of the intensity and pattern of CMF608 expression in bone tissue.

In sections of ectopic bones hybridization signal for CMF608 can be seen in some fibroblast-like cell either scattered within unmineralized connective tissue matrix or concentrated at the boundary between this tissue and osteoblasts of immature bone.

Pattern of expression of CMF608 gene revealed by in situ hybridization in bone and cartilage allows to speculate that its expression marks some skeletal tissue elements able to differentiate into two skeletal cell types - osteoblasts and chondrocytes. The terminal differentiation of these cells appears to be accompanied by down-regulation of CMF608 expression. The latter suggestion is supported by peculiar temporal pattern of CMF608 expression in primary cultures of osteogenic cells isolated from calvaria bones of rat fetuses. In these cultures expression was revealed by in situ hybridization in vast majority of cells after one and two weeks of incubation in vitro. Three and four weeks old cultures showing signs of ossification contain no expressing cells. Significantly, no hybridization signal was found on sections of multitissue block hybridized to CMF608 probe suggesting high specificity of this gene expression for the skeletal tissue in adult organism.

In situ hybridization study of embryonic sections demonstrated that at 12.5 dpc weak hybridization signal can be discerned in some mesenchymal cells in several locations throughout the embryonic body. The most prominent signal is found in the head: in loose mesenchymal tissue surrounding the olfactory epithelium and underlying the surface epithelium of nose tip. Other mesenchymal cells in the head also show hybridization signal: in non-cartilaginous part of basisphenoid bone primordium and in mesenchyme surrounding the dental laminae (tooth primordia) in the mandible.

In the trunk expression can be detected in less developed vertebrae primordia in thoraco-lumbar region. Hybridization signal here marks the condensed portion of

sclerotomes. Another area showing hybridization signal in the trunk is comprised by thin layer of mesenchymal cells in the anterior part of thoracic body wall.

At later stages of development -14.5 and 16.5 dpc probe CMF608 gave no hybridization signal. Thus, it appears that during embryonic development CMF608 gene is transiently expressed by at least some mesenchymal and skeleton-forming cells. This expression is down-regulated at later stages of development. More detailed study of late embryonic and postnatal stages of development will reveal the timing of appearance of CMF608 expressing cells in bone tissue.

#### Example/Result 4: CMF405 GENE EXPRESSION BY *IN SITU*

##### 10 HYBRIDIZATION

Pattern of expression of CMF405 gene was studied by in situ hybridization on sections of bones from ovariectomized and sham-operated rats. Female Wistar rats weighting 300-350 g were subjected to ovariectomy under general anesthesia. Control rats were operated in the same way but ovaries were not excised - sham operation.

15 Three weeks after operation rats were sacrificed and tibia were excised together with the knee joint. Bones were fixed for three days in 4% paraformaldehyde and then decalcified for four days in solution containing 5% formic acid and 10% formalin. Decalcified bones were postfixed in 10% formalin for three days and embedded into paraffin.

20 To study the pattern of expression of CMF405 gene in developing bone the model of ectopic bone formation was employed. Rat bone marrow cells were seeded into cylinders of demineralized bone matrix prepared from rat tibiae. Cylinders were implanted subcutaneously into adult rats. After three weeks rats were sacrificed and implants were decalcified and embedded into paraffin as described above for tibial bones.

25 The 6  $\mu$ m sections were prepared and subjected to in situ hybridization procedure. After hybridization sections were dipped into nuclear track emulsion and exposed for three weeks at 4°C. Autoradiographs were developed, stained with hematoxylin-eosin and studied under microscope using brightfield and darkfield illumination.

30 For the further assessment of cell and tissue specificity of CMF405 gene expression in situ hybridization study was performed on sections of multitissue block containing multiple samples of adult rat tissues. Developmental pattern of CMF405

expression was studied on sagittal sections of mouse embryos of 12.5, 14.5 and 16.5 days postconception (dpc) stages.

Bones: Hybridization signal for CMF405 gene is widely spread throughout different cell types on sections of long bones from sham-operated animals: cartilage,  
5 bone marrow and bone.

In the growth plate hybridization signal is concentrated in the transition zone from proliferating to hypertrophic cartilage so that most advanced proliferating chondrocytes and youngest hypertrophic chondrocytes display expression. Both young proliferating chondrocytes and most of mature hypertrophic chondrocytes do  
10 not show hybridization signal. Chondrocytes of articulate cartilage show no hybridization signal.

Some (but not all) hematopoietic cells within the bone marrow show clear hybridization signal. Poor morphology of decalcified section stained with hematoxylin-eosin does not allow identification of expressing cell types.

15 Within the bone tissue hybridization signal can be seen in osteoblasts localized in primary spongiosa and secondary spongiosa in metaphysis (cancellous bone). Osteoblasts covering the surface of marrow cavity and Volkmann's canals in diaphyseal (compact) bone also display hybridization signal. Flat bone lining cells and osteocytes are not expressed in any part of the bone.

20 Ovariectomy did not result in change of the intensity and pattern of CMF405 expression. In ectopic bone hybridization signal concentrates mainly in osteoblasts of immature bone. This signal is weak or absent from osteoblasts embedded into bone matrix.

The pattern of hybridization of CMF405 gene in adult skeletal tissues suggests  
25 that its expression is characteristic for osteogenic and chondrogenic cells at intermediate stage of their differentiation preceding intensive matrix calcification.

Tissue expression: The CMF405 probe was hybridized to multitissue block sections. The hybridization signal can be seen mainly in epithelial cells in many organs and tissues suggesting the wide expression of this gene in adult tissues.

30 The hybridization signal of varying intensity can be seen in epithelial lining of the digestive system. Weak hybridization signal can be seen in basal cells of stratified squamous epithelium of esophagus. Weak signal is displayed by surface epithelium of fundic stomach. In pyloric stomach strong hybridization signal is displayed by cells

lining mucosal pits and weaker signal - by surface epithelium. In thin intestine expressing cells are localized in crypts and glands while villous epithelium shows no hybridization signal. Similar pattern is observed in colonic epithelium: weak hybridization signal can be seen only in crypts and not in villi. This pattern of hybridization throughout different parts of the alimentary canal allows to suggest that expression of CMF405 in digestive system appears to be confined mainly to actively proliferating epithelial cells and transition of epithelial cells into non-proliferative compartment (like suprabasal layers of esophagus or villous epithelium of intestine) is accompanied by down-regulation of CMF405 expression.

- 10 In urogenital system weak and diffuse signal can be seen in medullar part of kidney. The weak hybridization signal is displayed by transitional epithelium of kidney calyx, ureter and bladder. Hybridization signal is seen also in basal cells within seminiferous tubules. Low resolution of microautoradiographs does not allow unequivocal identification of expressing cells as spermatogonia or Sertoli cells.

- 15 All epidermal layers of skin show hybridization signal. Strongly expressing cells are localized also in all layers of the hair follicle.

Strong and uniform hybridization signal can be seen on sections of lymphoid organs: thymus, spleen and lymph nodes.

- 20 Positive hybridization was obtained also on eye sections: strong signal is displayed by corneal epithelium. Retina shows weaker hybridization signal throughout all layers excluding ganglion cell layer.

No hybridization signal can be seen in brain.

- 25 *In situ* hybridization study of sections of 12.5, 14.5 and 16.5 dpc embryos revealed strong and practically uniform hybridization signal throughout bodies of 12.5 and 14.5 dpc embryos. This suggests that all cell types at these stages of development express CMF405 gene. By 16.5 dpc stage expression appears to decline in some cells so that pattern of expression approaches to that in adult tissue although some structures showing no expression in adults display hybridization signal.

**Example/Result 5: EXPRESSION OF CMF274 GENE BY *IN SITU* HYBRIDIZATION**

Hybridization of CMF274 probe to the sections of knee joint demonstrated wide expression throughout the bone, cartilage and bone marrow tissues.

- 5 Hybridization signal can be seen in hematopoietic bone marrow cells. Accumulation of signal in cells of eosinophilic lineage is clear. Poor morphology of decalcified sections does not allow identification of other expressing cell types although it is apparent that not all the myeloid elements are expressing.

- 10 Chondrocytes display weak hybridization signal throughout all zones of growth plate. No expression is detected in epiphyseal articulate cartilage.

- Within the bone tissue hybridization signal can be seen in osteoblasts localized in primary spongiosa and secondary spongiosa in metaphysis (cancellous bone). Periosteal and endosteal osteoblasts and osteoblasts of Volkmann's canals in diaphyseal (compact) bone also display hybridization signal. Flat lining cells and osteocytes are not expressing in any part of the bone. Ovariectomy did not result in change of the intensity and pattern of CMF274 expression in bone tissue.

The CMF274 probe was hybridized to multitissue block sections and to sections of 12.5, 14.5 and 16.5 dpc mouse embryos. No hybridization signal was found on these sections.

20 **Example/Result 6: CMF2-45 (SARP) GENE EXPRESSION BY *IN SITU* HYBRIDIZATION**

Hybridization signal for CMF2-45 (SARP) gene is found in different cell types on sections of long bones from sham-operated animals: cartilage, bone marrow and bone. In all expressing cell types the level of hybridization signal is rather low.

- 25 In growth plate hybridization signal marks proliferating chondrocytes while hypertrophic chondrocytes show little or no signal. Chondrocytes of articulate cartilage in epiphysis do not show hybridization signal.

- Hybridization signal in bone tissue proper marks osteoblasts located in all compartments of cancellous and compact bone: primary and secondary spongiosa, periosteum, endosteum and Volmann's canals. Bone lining cells and osteocytes show no hybridization signal.

Some (but not all) hematopoietic cells within the bone marrow show clear hybridization signal. Poor morphology of decalcified sections stained with hematoxylin-eosin does not allow identification of expressing cell types.

- In ectopic bone hybridization signal can be seen in osteoblasts of immature  
5 bone. Beside of osteoblasts single fibroblast-like cells scattered throughout the connective tissue also show hybridization signal. Some of these expressing fibroblast-like cells can be seen in close contact with osteoblasts.

**Example/Result 7: CMF2-224 (Rb13-6) EXPRESSION  
BY IN SITU HYBRIDIZATION**

- 10 *In situ* hybridization study of RB13-6 gene (GenBank Accession No.: Z47987) revealed expression of this gene in osteoblasts located in different compartments of long bone: in primary and secondary spongiosa of cancellous bone in metaphysis and within endosteum and Volkmann's canals of compact bone in diaphysis. Weak hybridization signal can be seen also in myeloid cells of bone marrow.  
15 immature bone developing within ectopic bone implants also display hybridization signal. Significantly, no hybridization signal was revealed on sections of long bones of ovariectomised rats and also on sections of ectopic bones implanted into ovariectomised rats. This result suggests that expression of RB13-6 in osteoblasts and bone marrow cells is estrogen-dependent. Further study will be needed to clarify the  
20 involvement of RB13-6 gene product into regulation of osteoblast function and development of osteoporosis.

- Additional *in situ* hybridization study on sections of multitissue block demonstrated expression of RB13-6 gene in distinct epithelial cell types and in lymphoid tissue. Expressing cells can be seen in epithelial lining of bronchi, villous  
25 (i.e. mature and non-proliferating) epithelium of thin intestine, in luminal and glandular epithelia of uterus, acinar and ductal epithelia of salivary glands. Very weak hybridization signal suggesting low level of expression was found in liver and in kidney. Liver expression appears to be uniform throughout hepatocytes. In kidney single expressing cells can be seen in thick ascending part of Henle's loop.

- 30 Within lymphoid tissue strongly expressing lymphocytes are concentrated in medullar zone of lymph nodes. Few expressing cells can be seen in spleen: in the perifollicular zone (the bordering area between the red and white pulp) and in perivascular aggregations of lymphocytes. No expressing cells were found in thymus.

Having thus described in detail preferred embodiments of the present invention, it is to be understood that the invention defined by the appended claims is not to be limited by particular details set forth in the above description as many  
5   apparent variations thereof are possible without departing from the spirit or scope thereof.

References

- Consensus Development Conference: Prophylaxis and Treatment of Osteoporosis: *Am J Med* 94:646-650, 1993
- 5 WHO Study Group: In *Assessment of Fracture Risk and Its Application to Screening for Postmenopausal Osteoporosis*. WHO Technical Report Series: 843, 1994
- Department of Health *Health and Personal Social Services Statistics for England, 1991 Edition*. HMSO, London, 1991, pp 10-11
- 10 Welsh Office *Health and Personal Social Services for Wales* HMSO. London, 1991, pp 2-3
- 15 Kanis JA, McCloskey EV: Epidemiology of vertebral osteoporosis. *Bone* 13:S1-S10, 1992
- Kanis JA, Pitt FA: Epidemiology of osteoporosis. *Bone* 13:S7-S15, 1992
- 20 Cummings SR et al: Epidemiology of osteoporosis and osteoporotic fractures. *Epidemiol Rev* 7:178-207, 1985
- Department of Health and Social Security Office of Population Censuses and Surveys *Cancer Statistics: Registrations*. Series MBT, 20, HMSO, London, 1987
- 25 Riggs BL, Melton LJ: Involutional osteoporosis. *N Engl J Med* 314:1676-1686, 1986
- Melton LJ, Eddy DM, Johnston CC: Screening for osteoporosis. *Ann Intern Med* 112:516-528, 1990
- 30 Department of Health and Social Security Office of Population Censuses and Surveys *Hospital In-Patient Enquiry Summary Tables*. Series MB4, 26, HMSO, London, 1985
- Nordin BEC, Need AG: How can we prevent osteoporosis? In *Osteoporosis 1987* (Christiansen C, Riis BJ, eds.). Osteopress ApS, Copenhagen, 1987, pp 1204-1210
- 35 United Nations *Interpolated National Populations by Age and Sex: 1950-2025 (1992 Revision)*. United Nations Population Division, United Nations, New York, 1992
- 40 Cooper C, Campion G, Melton LJ: Hip fractures in the elderly: A world-wide projection. *Osteoporosis Int* 2:285-289, 1992
- Parfitt AM: Quantum concept of bone remodeling and turnover: Implications for the pathogenesis of osteoporosis. *Calcif Tissue Int* 28: 1-5, 1979
- 45 Parfitt AM: Bone remodeling in the pathogenesis of osteoporosis. *Res Staff Physician, December:60-72*, 1981



- Parfitt AM: Bone remodeling: Relationship to the amount and structure of bone, and the pathogenesis and prevention of fractures. In *Osteoporosis: Etiology, Diagnosis, and Management* (Riggs BL, Melton LJ, eds.). Raven Press, New York, 1988, pp 74-93
- 5 Kanis JA *Pathophysiology and Treatment of Paget's Disease of Bone*. Martin Dunitz, London, 1991, pp 12-40
- 10 Smith DM et al: Genetic factors in determining bone mass. *J Clin Invest* 52:2800-2808, 1973
- Pocock NA et al: Genetic determinants of bone mass in adults: A twin study. *J Clin Invest* 80:706-710, 1987
- 15 Matkovic V et al: Factors that influence peak bone mass formation: A study of calcium balance and the inheritance of bone mass in adolescent females. *Am J Clin Nutr* 52:878-888, 1990
- 20 Cohn SH et al: Comparative skeletal mass and radial bone mineral content in black and white women. *Metabolism* 26:171-178, 1977
- Selby PL: Endocrinology and osteoporosis. In *HRT and Osteoporosis* (Drife JO, Studd JWW, eds.). Springer Verlag, London, 1990, pp 105-114
- 25 Stevenson JC et al: Exercise and the skeleton. In *Osteoporosis 1990* (Smith R, ed.). Royal College of Physicians of London, London, 1990, pp 119-124
- Griffin J: *Osteoporosis and the Risk of Fracture*. Office of Health Economics, London, 1990, p 48
- 30 Mundy GR *Calcium Homeostasis: Hypercalcemia and Hypocalcemia*, 2nd ed. Martin Dunitz, London, 1990, pp 1, 32, 51-52, 179, 200
- Christiansen C, Riis B *The Silent Epidemic: Postmenopausal Osteoporosis*. Christiansen and Riis, Denmark, 1990, pp 23, 66-67, 70-73
- Christiansen C, Rodbro P, Tjellesen L: Serum alkaline phosphatase during hormone treatment in early postmenopausal women: A model for establishing optimal prophylaxis and treatment in postmenopausal osteoporosis. *Acta Med Scand* 216:11-17, 1984
- 40 Gallagher JC et al: Intestinal calcium absorption and serum vitamin D metabolites in normal subjects and osteoporotic patients: Effect of age and dietary calcium. *J Clin Invest* 64:729-736, 1979
- 45 Heaney RP: Nutritional factors in bone health. In *Osteoporosis: Etiology, Diagnosis, and Management* (Riggs BL, Melton LJ, eds.). Raven Press, New York, 1988, pp 359-372

- Heaney RP: Effect of calcium on skeletal development, bone loss, and risk of fractures. *Am J Med* 91 (suppl 5B):5B-23S-5B-29S, 1991
- 5 Delmas PD et al: Increase in serum bone gamma-carboxyglutamic acid protein with aging in women: Implications for the mechanism of age-related bone loss. *J Clin Invest* 71:1316-1321, 1983
- Drinkwater BL et al: Bone mineral content of amenorrheic and eumenorrheic athletes. *N Engl J Med* 311:277-281, 1984
- 10 Silverman SL: The clinical consequences of vertebral compression fracture. *Bone* 13:S27-S31, 1992
- 15 Adami S et al: The radiological assessment of vertebral osteoporosis. *Bone* 13:S33-S36, 1992
- Riggs BL: Practical management of the patient with osteoporosis. In *Osteoporosis: Etiology, Diagnosis, and Management* (Riggs BL, Melton LJ, eds.). Raven Press, New York, 1988, pp 481-490
- 20 Kanis JA et al: Screening techniques in the evaluation of osteoporosis. In *HRT and Osteoporosis* (Drife JO, Studd JWW, eds.). Springer Verlag, London, 1990, pp 135-147
- 25 Genant HK, Faulkner KG, Gluer C-C: Measurement of bone mineral density: Current status. *Am J Med (suppl 5B)* 91:5B-49S-5B-53S, 1991
- Sartoris DJ, Resnick D: Digital radiography may spark renewal of bone densitometry. *Diagnostic Imaging*, January:145-150, 1988
- 30 Fogelman I, Rodin A: The measurement of bone density. In *HRT and Osteoporosis* (Drife JO, Studd JWW, eds.). Springer Verlag, London, 1990, pp 119-133
- 35 Mazess RB et al: Dual-energy X-ray absorptiometry for total-body and regional bone-mineral and soft-tissue composition. *Am J Clin Nutr* 51:1106-1112, 1990
- Johnston CC et al: Clinical indications for bone mass measurements. *J Bone Min Res* 4 (suppl 2): 1-29, 1989
- 40 Delmas PD: Clinical use of biochemical markers of bone remodeling in osteoporosis. In *Osteoporosis 1990* (Christiansen C, Overgaard K, eds.). Osteopress ApS, Copenhagen, 1990, pp 450-458
- 45 Uebelhart D et al: Urinary excretion of pyridinium crosslinks: A new marker of bone resorption in metabolic bone disease. *Bone Mineral* 8:87-96, 1990
- Consensus Development Conference: Prophylaxis and Treatment of Osteoporosis: *Am J Med* 90:107-110, 1991

- Delmas PD: Biochemical markers for the assessment of bone turnover. In *Osteoporosis: Etiology, Diagnosis, and Management*. Rev. ed. (Riggs BL, Melton LJ, eds.). Raven Press, New York, 1995, pp 319-333
- 5 Uebelhart D et al: Effect of menopause and hormone replacement therapy on the urinary excretion of pyridinium cross-links. *J Clin Endocrinol Metab* 72:367-373, 1991
- 10 Christiansen C, Riis BJ, Rodbro P: Screening procedure for women at risk of developing postmenopausal osteoporosis. *Osteoporosis Int* 1:35-40, 1990
- Sinaki M: Exercise and physical therapy. In *Osteoporosis: Etiology, Diagnosis, and Management* (Riggs BL, Melton LJ, eds.). Raven Press, New York, 1988, pp 457-479
- 15 Davee AM et al: Exercise patterns and trabecular bone density in college women. *J Bone Min Res* 5:245-250, 1990
- Gleeson PB et al: Effects of weight lifting on bone mineral density in premenopausal women. *J Bone Min Res* 5:153-158, 1990
- 20 Riggs BL, Melton LJ: The prevention and treatment of osteoporosis. *N Engl J Med* 327:620-627, 1992
- Riggs BL: Treatment of osteoporosis with sodium fluoride or parathyroid hormone. *Am J Med* (suppl 5B) 91:5B-37S-5B-41S, 1991
- 25 Kanis JA et al: Osteoporosis: Causes and therapeutic implications. In *Osteoporosis 1990* (Smith R, ed.). Royal College of Physicians of London, London, 1990, pp 45-56
- 30 Rogers HJ, Spector AG, Trousseau JR: In *A Textbook of Clinical Pharmacology*. Hodder and Stoughton, London, 1981, p 583
- Eriksen EF et al: Evidence of estrogen receptors in normal human osteoblast-like cells. *Science* 241 :84-86, 1988
- 35 Lindsay R, Cosman F: Estradiol in prevention of osteoporosis. In *Osteoporosis 1990* (Christiansen C, Overgaard K, eds.). Osteopress ApS, Copenhagen, 1990, pp 2070-2075
- 40 Lindsay R: Estrogens, bone mass, and osteoporotic fracture. *Am J Med* 91 (suppl 5B):5B10S-5B-13S, 1991
- Lindsay R, Tohme JF: Estrogen treatment of patients with established postmenopausal osteoporosis. *Obstet Gynecol* 76:290-295, 1990
- 45 Christiansen C, Riis BJ: 17 $\beta$ -estradiol and continuous norethisterone: A unique treatment for established osteoporosis in elderly women. *J Clin Endocrinol Metab* 71:836-841, 1990

- Lufkin EG et al: Treatment of postmenopausal osteoporosis with transdermal estrogen. *Ann intern Med* 117:1-9, 1992
- Segre G, Bruni G, Dal Pra P: Calcitonin pharmacokinetics. In *Calcitonin* (Pecile A, ed.). Elsevier, New York, 1985, pp 99-107
- Nicholson GC et al: Abundant calcitonin receptors in isolated rat osteoclasts. *J Clin Invest* 78:355-360, 1986
- 10 Baron R, Vignery A: Behaviour of osteoclasts during a rapid change in their number induced by high doses of parathyroid hormone or calcitonin in intact rats. *Metab Bone Dis Rel Res* 2:339-343, 1981
- Overgaard K et al: Effect of calcitonin given intranasally on early postmenopausal bone loss. *Br Med J* 299:477-479, 1989
- 15 Rasmussen H et al: Effect of combined therapy with phosphate and calcitonin on bone volume in osteoporosis. *Metab Bone Dis Rel Res* 2:107-111, 1980
- 20 Gruber HE et al: Long-term calcitonin therapy in postmenopausal osteoporosis. *Metabolism* 33:295-303, 1984
- Christiansen C et al: Intranasal calcitonin in the treatment of postmenopausal osteoporosis. In *Osteoporosis 1987* (Christiansen C, Riis BJ, eds.). Osteopress ApS, Copenhagen, 1987, pp 853-856
- 25 Overgaard K et al: Nasal calcitonin for treatment of established osteoporosis. *Clin Endocrinol* 30: 435-442, 1989
- 30 Overgaard K et al: Discontinuous calcitonin treatment of established osteoporosis: Effects of withdrawal of treatment. *Am J Med* 89:1-6, 1990
- Fleisch H: Bisphosphonates: Pharmacology and use in the treatment of tumour-induced hypercalcemic and metastatic bone disease. *Drugs* 42:919-944, 1991
- 35 Watts NB et al: Intermittent cyclical etidronate treatment of postmenopausal osteoporosis. *N Engl J Med* 323:73-79, 1990
- Filipponi P et al: Cyclical clodronate is effective in preventing postmenopausal bone loss: A comparative study with transcutaneous hormone replacement therapy. *J Bone Min Res* 10:697-703, 1995
- 40 Chesnut CH et al: Alendronate treatment of the postmenopausal osteoporotic woman: Effect of multiple dosages on bone mass and bone remodeling. *Am J Med* 99: 144-152, 1995
- 45 Garnero P et al: Comparison of new biochemical markers of bone turnover in late postmenopausal osteoporotic women in response to alendronate treatment. *J Clin Endocrinol Metab* 79(6):1693-1700, 1994

- Francis RM: The calcium controversy. In *Osteoporosis 1990* (Smith R, ed.). Royal College of Physicians of London, London, 1990, pp 125-133
- 5 Johnston CC et al: Calcium supplementation and increases in bone mineral density in children. *N Engl J Med* 327:82-87, 1992
- Baran D et al: Dietary modification with dairy products for preventing vertebral bone loss in premenopausal women: A three-year prospective study. *J Clin Endocrinol Metab* 70:264-270, 1989
- 10 Dawson-Hughes B et al: A controlled trial of the effect of calcium supplementation on bone density in postmenopausal women. *N Engl J Med* 323:878-883, 1990
- 15 Chapuy MC et al: Vitamin D3 and calcium to prevent hip fractures in elderly women. *N Engl J Med* 327:1637-1642, 1992
- Hassager C et al: Changes in soft tissue body composition and plasma lipid metabolism during nandrolone decanoate therapy in postmenopausal osteoporotic women. *Metabolism* 38:238-242, 1989
- 20 Johansen JS et al: Treatment of postmenopausal osteoporosis: Is the anabolic steroid nandrolone decanoate a candidate? *Bone Mineral* 6:77-86, 1989
- 25 Riggs BL: Treatment of osteoporosis with sodium fluoride: An appraisal. In *Bone and Mineral Research, Annual 2* (Peck WA, ed.). Elsevier, New York, 1983, pp 366-393
- Delmas PD: Treatment of vertebral osteoporosis with disodium monfluorophosphate: Comparison with sodium fluoride. *J Bone Mineral Res* 5(suppl 1):S143-S147, 1990
- 30 Riggs BL et al: Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. *N Engl J Med* 322:802-809, 1990
- 35 Mamellet N et al: Risk-benefit ratio of sodium fluoride treatment in primary vertebral osteoporosis. *Lancet* ii:361-365, 1988
- Tietz NW, ed. *Textbook of Clinical Chemistry*. WB Saunders, Philadelphia, 1986, pp 1322-1323
- 40 Russell RGG: Bone cell biology: The role of cytokines and other mediators. In *Osteoporosis 1990* (Smith R, ed.). Royal College of Physicians of London, London, 1990, pp 9-33
- Tilyard MW et al: Treatment of postmenopausal osteoporosis with calcitriol or calcium. *N Engl J Med* 326:357-362, 1992
- 45 Meunier PJ et al: Can we stop bone loss and prevent hip fractures in the elderly? *Osteoporosis Int* 4(suppl 1):S71-S76, 1994

- Chapuy MC et al: Effect of calcium and cholecalciferol treatment for three years on hip fractures in elderly women. *Br Med J* 308:1081-1089, 1994
- 5 Neer R et al: Treatment of postmenopausal osteoporosis with daily parathyroid hormone plus calcitriol. In *Osteoporosis 1990* (Christiansen C, Overgaard K, eds.). Osteopress ApS, Copenhagen, 1990, pp 1314-1317
- Reid DM: Corticosteroid-induced osteoporosis. In *Osteoporosis 1990* (Smith R, ed.). Royal College of Physicians of London, London, 1990, pp 99-117
- 10 Lukert BP, Raisz LG: Glucocorticoid-induced osteoporosis: Pathogenesis and management. *Ann Intern Med* 112:352-364, 1990
- Reid IR et al: Two-year follow-up of bisphosphonate (APD) treatment in steroid-induced osteoporosis. *Lancet* ii: 1144, 1988
- 15 Ringe JD, Welzel D: Salmon calcitonin in the therapy of corticoid-induced osteoporosis. *Eur J Clin Pharmacol* 33:35-39, 1987
- 20 Rosen, CJ: Beyond one gene-one disease: Alternative strategies for deciphering genetic determinants of osteoporosis. *Calcif Tissue Int* 60(3):225-28, 1997.
- Keen et al.: Genetic factors in osteoporosis. What are the implications for prevention and treatment? *Drugs Aging (New Zealand)* 11(5):333-7, 1997.
- 25 Eisman: Vitamin D receptor gene variants: Implications for therapy. *Curr Opin Genet Dev* 6(3): 361-65, 1996.
- Ralston: The genetics of osteoporosis. *QJM - Monthly Journal of the Association of Physicians (United Kingdom)* 90/4, pp 247-51, 1997.
- 30 Cole et al.: Osteoporosis And the Interface Between Nutrition And Genetics. INABIS '98 - 5<sup>th</sup> Internet World Congress on Biomedical Sciences at McMaster University, Canada, Dec 7-16. <http://www.mcmaster.ca/inabis98/atkinson/cole0195/index.html>
- 35 Tokita et al.: Vitamin D receptor alleles, bone mineral density and turnover in premenopausal Japanese women. *J Bone Miner Res* 11(7): 1003-09, 1996.
- Jouanny et al.: Environmental and genetic factors affecting bone mass. Similarity of bone density among members of healthy families. *Arthritis Rheum* 38(1):61-67, 1995.
- 40 Garnero et al.: Genetic influence on bone turnover in postmenopausal twins. *J Clin Endocrinol Metab* 81(1):140-46, 1996.
- 45 Johnston et al.: Pathogenesis of osteoporosis. *Bone* 17(2 Suppl):19S-22S, 1995.
- Fujita: Vitamin D in the treatment of osteoporosis revisited. *Proc Soc Exp Biol Med* 212(2):110-5, 1996.

- Gong et al.: Osteoporosis-pseudoglioma syndrome, a disorder affecting skeletal strength and vision, is assigned to chromosome region 11q12-13. *Am J Hum Genet* 59(1):146-51, 1996.
- 5 Houston et al.: Vitamin D receptor polymorphism, bone mineral density, and osteoporotic vertebral fracture: studies in a UK population. *Bone* 18(3):249-52, 1996.
- Cummings: Treatable and untreatable risk factors for hip fracture. *Bone* 18(3 Suppl):165S-167S, 1996.
- 10 Wasnich: Vertebral fracture epidemiology. *Bone* 18(3 Suppl):179S-183S, 1996.
- Riggs et al.: The contribution of vitamin D receptor gene alleles to the determination of bone mineral density in normal and osteoporotic women. *J Bone Miner Res* 15 10(6):991-6, 1995.
- Wark: Osteoporotic fractures: background and prevention strategies. *Maturitas* (Ireland) 23(2):193-207, 1996.
- 20 Fleet et al.: The BsmI vitamin D receptor restriction fragment length polymorphism (BB) predicts low bone density in premenopausal black and white women. *J Bone Miner Res* 10(6):985-90, 1995.
- Krall et al.: Vitamin D receptor alleles and rates of bone loss: influences of years since menopause and calcium intake. *J Bone Miner Res* 10(6):978-84, 1995.
- 25 Livshits et al.: Population biology of human aging: segregation analysis of bone age characteristics. *Hum biol* 68(4):540-54, 1996.
- 30 Dalglish: The human type I collagen mutation database. *Nucleic Acids Res* 25(1):181-7, 1997 (citing <http://www.le.ac.uk/depts/ge/collagen/collagen.html>)
- Lonzer et al.: Effects of heredity, age, weight, puberty, activity and calcium intake on bone mineral density in children. *Clin Pediatr* 35(4):185-9, 1996.
- 35 Pereira et al.: Bone fragility in transgenic mice expressing a mutated gene for type I procollagen (COL1A1) parallels the age-dependent phenotype of human osteogenesis imperfecta. *J Bone Miner Res* 10(12):1837-43, 1995.
- 40 Sano et al.: Association of estrogen receptor dinucleotide repeat polymorphism with osteoporosis. *Biochem Biophys Res Commun* 217(1):378-83, 1995.
- Burke and Olson, "Preparation of Clone Libraries in Yeast Artificial-Chromosome Vectors" in *Methods in Enzymology*, Vol. 194, "Guide to Yeast Genetics and Molecular Biology", eds. C. Guthrie and G. Fink, Academic Press, Inc., Chap. 17, pp. 251-270 (1991).
- 45 Capecchi, "Altering the genome by homologous recombination" *Science* 244:1288-1292 (1989).

- Davies et al., "Targeted alterations in yeast artificial chromosomes for inter-species gene transfer", *Nucleic Acids Research*, Vol. 20, No. 11, pp. 2693-2698 (1992).
- 5 Dickinson et al., "High frequency gene targeting using insertional vectors", *Human Molecular Genetics*, Vol. 2, No. 8, pp. 1299-1302 (1993).
- Duff and Lincoln, "Insertion of a pathogenic mutation into a yeast artificial chromosome containing the human APP gene and expression in ES cells", *Research*  
10 *Advances in Alzheimer's Disease and Related Disorders*, 1995.
- Huxley et al., "The human HPRT gene on a yeast artificial chromosome is functional when transferred to mouse cells by cell fusion", *Genomics*, 9:742-750 (1991).
- 15 Jakobovits et al., "Germ-line transmission and expression of a human-derived yeast artificial chromosome", *Nature*, Vol. 362, pp. 255-261 (1993).
- Lamb et al., "Introduction and expression of the 400 kilobase precursor amyloid protein gene in transgenic mice", *Nature Genetics*, Vol. 5, pp. 22-29 (1993).
- 20 Pearson and Choi, Expression of the human  $\alpha$ -amyloid precursor protein gene from a yeast artificial chromosome in transgenic mice. *Proc. Natl. Acad. Sci. USA*, 1993. 90:1057882.
- 25 Rothstein, "Targeting, disruption, replacement, and allele re-scue: integrative DNA transformation in yeast" in *Methods in Enzymology*, Vol. 194, "Guide to Yeast Genetics and Molecular Biology", eds. C. Guthrie and G. Fink, Academic Press, Inc., Chap. 19, pp. 281-301 (1991).
- 30 Schedl et al., "A yeast artificial chromosome covering the tyrosinase gene confers copy number-dependent expression in transgenic mice", *Nature*, Vol. 362, pp. 258-261 (1993).
- 35 Strauss et al., "Germ line transmission of a yeast artificial chromosome spanning the murine  $\alpha 1(I)$  collagen locus", *Science*, Vol. 259, pp. 1904-1907 (1993).
- Gilboa, E, Eglitis, MA, Kantoff, PW, Anderson, WF: Transfer and expression of cloned genes using retroviral vectors. *BioTechniques* 4(6):S04-512, 1986.
- 40 Cregg JM, Vedvick TS, Raschke WC: Recent Advances in the Expression of Foreign Genes in *Pichia pastoris*, *Bio/Technology* 11:905-910, 1993
- Culver, 1998. Site-Directed recombination for repair of mutations in the human ADA gene. (Abstract) *Antisense DNA & RNA based therapeutics*, February, 1998, Coronado, CA.
- 45 Huston et al, 1991 "Protein engineering of single-chain Fv analogs and fusion proteins" in *Methods in Enzymology* (ii Langone, ed.; Academic Press, New York, NY) 203:46-88.



- Johnson and Bird, 1991 "Construction of single-chain Fv derivatives of monoclonal antibodies and their production in *Escherichia coli* in *Methods in Enzymology* (JJ Langone, ed.; Academic Press, New York, NY) 203:88-99.
- 5 Mernaugh and Mernaugh, 1995 "An overview of phage-displayed recombinant antibodies" in *Molecular Methods In Plant Pathology* (RP Singh and US Singh, eds.; CRC Press Inc., Boca Raton, FL) pp. 359-365.
- 10 Aiello et al., "Identification of multiple genes in bovine retinal pericytes altered by exposure to elevated levels of glucose by using mRNA differential display" *Proc. Natl. Acad. Sci. USA* Vol. 91, pp. 6231-6235 (1994).
- 15 Bauer et al., "Identification of differentially expressed mRNA species by an improved display technique (DDRT-PCR)" *Nucleic Acids Research* Vol. 21, No. 18 (1993).
- Bhanicha and Ven Murthy, "Characterization of Polysomes and Polysomal mRNAs by Sucrose Density Gradient Centrifugation Followed by Immobilization in Polyacrylamide Gel Matrix" *Methods in Enzymology* Vol. 216, pp. 168-179 (1992).
- 20 Braun et al., "Identification of Target Genes for the Ewing's Sarcoma EWS/FLI Fusion Protein by Representational Difference Analysis" *Molecular and Cellular Biology* Vol. 15, No. 8, pp. 4623-4630 (1995).
- 25 Cold Springs Harbor Laboratory Press, Cold Spring Harbor, New York, 1989
- Davis et al., "Expression of a single transfected cDNA converts fibroblasts to myoblasts." *Cell* 51:9871000, 1987.
- 30 Diatchenko et al., "Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries" *Proc. Nad Acad. Sci.*, Vol. 93, pp. 6025-6030 (1996).
- Ehrenfeld, "Initiation of Translation by Picornavirus RNAs", *Translational Control* Cold Spring Harbor Laboratory Press, pp. 549-573, 1996.
- 35 Hadman et al., "Modification to the differential display technique reduce background and increase sensitivity" *Analytical Biochemistry* 226:383-386 (1995).
- 40 Hanauske-Abel et al., "Detection of a sub-set of polysomal mRNAs associated with modulation of hypusine formation at the G1-S boundary. Proposal of a role for EIF-5A in onset of DNA replication." *FEBS Letters* 386 pp. 92-98 (1995).
- Hirama et al., "Direct Purification of Polyadenylated RNAs from Isolated Polysome Fractions" *Analytical Biochemistry* 155, pp. 385-390 (1986).
- 45 Hubank and Schatz, "Identifying differences in mRNA expression by representational difference analysis of cDNA" *Nucleic Acids Research*, Vol. 22, No. 25, p. 5640-5648 (1994). Jefferies et al., "Elongation Fraction-la mRNA Is Selectively Translated

- following Mitogenic StimulatioW' The Journal of Biological Chemistry Vol. 269, No. 6, pp. 4367-4372 (1994).
- 5 Liang and Pardee, "Differential Display of Eukaryotic Messenger RNA by Means of the Polymerase Chain Reaction" ~S~cience, Vol. 257, pp. 967-971 (1992).
- Liang et al., "Distribution and cloning of Eukaryotic mRNAs by means of differential display: refinements and optimization" Nucleic Acids Research Vol. 21, No. 14, pp. 3269-3275 (1993).
- 10 Liang and Pardee, "Recent advances in differential display" Current Opinion in Immunology, 7:274-280 (1995).
- Linskens et al., Cataloging altered gene expression in young and senescent cells using enhanced differential display" Nuc. Ac. Res. 23: 3244-3251 (1995).
- 15 Lisitsyn and Wigler, "Cloning the Differences Between Two Complex Genomes" Science Vol. 259, pp. 946-951 (1993).
- 20 Mach et al., "Isolation of a cDNA Clone Encoding S-Adenosylmethionine Decarboxylase" The Annual of Biological Chemistry Vol. 261, No. 25, pp. 11697-11703 (1986).
- 25 Macejak et al., "Internal inition of translation mediated by the 5' leader of a cellular mRNA" Nature, Vol. 353, pp. 990-94 (1991).
- Mechler, "Isolation of messenger RNA from Membrane-Bound Polysomes" Methods in Enzymology Vol. 152, pp. 241-253 (1987).
- 30 Menaker et al., "A Method for the Isolation of Rat Submandibular Salivary Gland Polysomes on Linear Sucrose Density Gradients" Analytical Biochemistry 57, pp. 325-335 (1974).
- Meyuhas et al., "Translational Control of Ribosomal Protein mRNAs in Eukaryotes" Translational Control,, pp. 363-388 (1996).
- 35 Mountford et al., "Internal ribosome entry sites and dicistronic RNAs in mammalian transgenesis" TIG, Vol. 11. No. 5, pp. 179-184 (1995).
- 40 Ogishima, et al., "Fractionation of Mammalian Tissue mRNAs by High-Performance Gel Filtration Chromatography" Analytical Biochemistry 138, pp. 309-313 (1984).
- Oh et al., "Gene regulation: translational initiation by internal ribosome binding" Current Opinion in Genetics and development pp. 295-300 (1993).
- 45 Pelletier et al., "Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA" Nature Vol. 334, pp. 320-325 (1988).
- Schena et al., "Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray" Science Vol. 270, pp. 467-470 (1995).

- Shen et al., "Identification of the Human Prostatic Carcinoma Oncogene PTI-1 by Rapid Expression Cloning and Differential RNA Display" *Proc. Natl. Acad. Sci. USA* Vol. 92, pp. 6778-6782 (1995).
- 5 Vagner et al. "Alternative Translation of Human Fibroblast Growth Factor 2 mRNA Occurs by Internal Entry of Ribosomes" *Molecular and Cellular Biology*, Vol. 15, No. 1, pp. 35-44 (1995).
- Welsh et al., "Arbitrary primed PCR fingerprinting of RNA", *Nuc. Ac. Res.*  
10 20:4965-4970 (1992).
- Zhao et al., "New primer strategy improves precision of differential display" *Biotechniques* 18: 842-850 (1995).
- 15 Sperling et al., "Abundant Nuclear Ribonucleoprotein Form of CAD RNA," *Mol Cell Biol.* 1985 Mar;5(3):569-75.
- Dimitrijevic & Wgner "A search for novel, naturally occurring antisense RNA systems"
- 20 Murphy, P.R. and Knee R.S. (1994) Identification and characterization of an antisense RNA transcript (gfg) from the human basic fibroblast growth factor gene. *Mol. Endo* Vol. 8 no. 7 852-859.
- 25 Adelman, J.P. Bond, C.T., Douglass, J. and Herbert, E. (1987) Two mammalian genes transcribed from opposite strands of the same DNA locus. *Science* 235:1514:1313.
- Kimelman, Du Kirschner, M.W. (1989) An antisense m-RNA directs the covalent modification of the transcript encoding pibroblast growth factor in *Xenopus* oocytes  
30 *Cell*, 59:687-696.
- Lazer, M.A. Hodin, R.A., Darling D.S., Chin, ww. (1989) A novel member of the thyroid/steroid hormone receptor family is encoded by the opposite strand of the rat c-*erbA* a transcriptional unit. *Mol. Cell. Biol.* 9:1128-1136.
- 35 Miyajima, N., Horiuchi, R., Shibuya, YI, Fuhushige, S., Mutsubara, K., Toyoshimak, Yamamoto, T (1989) Tow *erba* homologs encoding proteins with different T3 binding capacities are transcribed from opposite DNA strands of the same genetic locus. *Cell*. 57:31-39.
- 40 Kindy, M., McCormack, J. Buckler, A., Levine, R., Sanenshein G. (1987) Independent regulation of the transcription of the two strands of the c-myc gene. *Mol. Cell. Biol.* 7:2857-2562.
- 45 Celano, P., Berchtold, C., Kizer, D., Weararatna, F., Nelkin, B., Baylin, S., Casero, R. (1992) Characterization of an edogenous RNA transcript with homology to the antisense strand of the human c-myc gene. *J. Biol. Chem.* 267:15092-15096.

Armstrong, B.C., Krystal, G.W. (1992) Isolation and characterization of complementary DNA per N-cym, a gene encoded by the DNA strand opposite to N-myc Cell Growth Differ. 3:385-390.

5

Taylor, E.R., Seleiro, E.A.P. Brickell, P.M. (1991) Identification of antisense transcripts of the chicken insulin-like growth factor II gene. J. Mol. Endocrial 7:145

10

Mitchell, P.J., Carothers, M., Han, J.H., Harding, J.D., Kas, E. Vendia, L., Chasin, L.A. (1986) Multiple transcription start sites, Dnase I-hypersensitive sites, and an opposite strand exon in the 5' region of the CHO dnfr gene. Mol. Cell. Biol. 6:425-

- A search for novel, naturally occurring antisense RNA systems, N. Dimitrijevic,  
5 E.G.H. Wagner, Dept. of Microbiology, SLU (Swedish University of Agricultural  
Sciences), Genetic Center, P.O. Box 7025, S-75007 Uppsala, Sweden

Antisense RNA control of gene expression has been demonstrated in many  
bacteria, whereas only few cases are known in eukaryotes. All antisense RNAs  
10 identified to date have been found fortuitously. Therefore, our goal is to develop a  
novel strategy for targeted identification of naturally occurring antisense systems. The  
approach is based on the complementarity between antisense and target RNA over a  
significantly long stretch of nucleotides. The method used is briefly described here.  
Total cellular RNA is extracted. Part of the RNA pool is immobilized on a  
15 membrane, another part converted into cDNA after ligation of oligodeoxynucleotides  
to the 3' ends. The use of biotinylated, complementary oligos for cDNA synthesis  
allows immobilization of a "minus" strand to streptavidin-coated magnetic beads. A  
second set of oligos is ligated to the cDNA at the previous 5' end of the RNA. Plus  
strands are eluted from the bound strands and hybridized to the membrane-bound  
20 RNA. Since the cDNA strand used has the same polarity of the RNAs, only cDNA  
sequences that can bind to complementary RNAs should be retained. PCR  
amplification and subsequent cloning of PCR-fragments is followed by sequence  
analysis. To test whether cloned sequences are correctly identified, probes are  
generated in sense and antisense direction. Positive clones will be structurally and  
25 functionally characterized. In order to work out this method, we started using a  
bacterial strain (*Escherichia coli*), containing plasmid R1 that regulates its copy  
number by antisense RNA. Previous work has identified both antisense (CopA) and  
target RNA (CopT) of R1 intracellularly. This procedure, if feasible, will then be  
used to screen for antisense RNA systems in other organisms.

**WHAT IS CLAIMED IS**

1. A method or process for identifying genes whose expression is responsive to a specific cue or cues including the steps of:
  - (a) applying a cue to an organism or tissue or cells;
  - 5 (b) isolating specific cellular fractions from the tissues or cells subjected to the cue;
  - (c) extracting the mRNA from the cellular fractions; and
  - (d) differentially analyzing the mRNA samples in comparison with control samples not subjected to the cue to identify genes that have responded to the cue;
  - 10 wherein the tissue or cells comprises bone cells that retain being bone cells in a culture.
2. A method as set forth in claim 1, wherein genes are identified at the translation level; genes regulated at the transcription level; genes regulated by RNA stability; genes regulated by mRNA transport rate between the nucleus and  
15 cytoplasm; genes regulated by differential splicing; and genes regulated by antisense RNA.
3. The method of claim 1 wherein the cue comprises a stress inducing element which comprises mechanical stress.
4. The method of claim 1 wherein the bone cells comprise calvaria cells.
- 20 5. A method for determining risk of developing a physiological or disease state based upon presence or increase from normal cells or absence or decrease from normal cells of mRNA or protein from a gene shown to be up regulated or down regulated in a mammal by a method of claim 1 comprising:
  - (a) determining the level or status of mRNA in cells of said mammal; and/or
  - 25 (b) determining the level or status of corresponding protein in cells of said mammal; and
  - (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a  
30 transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining a

presence or increase from normal cells, or an absence or decrease from normal cells, of mRNA or protein and thus risk of developing a physiological or disease state.

6. A method for testing a medicament for or a gene therapy approach to a physiological or disease state or other factors causing or contributing thereto or to symptoms thereof based on absence or decrease from normal cells or presence or increase from normal cells of mRNA or protein of identified genes comprising a method as claimed in claim 5 additionally comprising: (a') administering the medicament or the gene therapy; and from the comparing determining an absence or decrease from normal cells or presence or increase from normal cells of the relevant mRNA or protein and thus efficacy of the medicament or the gene therapy.

7. A method for treating, preventing or controlling a physiological or disease state comprising a method as claimed in claim 5 and additionally comprising administering a medicament or treatment therefor or for a cause thereof or a symptom thereof.

8. The method of claim 6 wherein the medicament or treatment comprises the protein, a functional portion thereof, a vector expressing the protein or a functional portion thereof, or an inhibitor of the protein or of a functional portion thereof, or an inhibitor of a nucleic acid encoding the protein or a functional portion thereof.

9. The method of claim 5 further comprising:

(d) determining the level or status of a second gene mRNA in cells of said mammal; and/or

(e) determining the level or status of protein expressed by a second gene product in cells of said mammal; and

(f) comparing said level or status of that mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining risk.

10. The method of claim 9 wherein steps (a) and/or (b) and optionally (d) and/or (e) are carried out *in vivo* and/or steps (a) and/or (b) and optionally (d) and/or (e) are carried out *in vitro*.

11. The method of claim 9 wherein the determination in step (a) and  
5 optionally in step (d) is effected by employing
- (i) a nucleic acid sequence corresponding to at least a part of the gene encoding at least part of the protein and optionally a second nucleic acid sequence corresponding to at least a part of the second gene encoding at least part of the second protein;
  - (ii) a nucleic acid sequence complementary to the nucleic acid sequence(s) of (i);
  - 10 or
  - (iii) a primer or a primer pair hybridizing to the nucleic acid sequence(s) of (i) or (ii).

12. The method of claim 9 wherein the determination in step (b) and optionally of step (e) is effected by employing an antibody or a fragment thereof that  
15 specifically binds to the protein and optionally by employing a second antibody or a fragment thereof which specifically binds to the second protein.

13. A gene identification method comprising: preparation of probes from a model system; analysis of DNA chip hybridization; sequencing of clones showing differential expression; and optionally full-length cloning of clones of interest;  
20 wherein the model system comprises bone cells which retain their characteristic thereof in cultures which have mechanical stress or a lack thereof applied thereto.

14. The method of claim 13 wherein the bone cells comprise a calvaria primary culture.

15. A method for determining risk of developing osteoporosis or low or  
25 high bone density or other factors causing or contributing to osteoporosis or other conditions involving mechanical stress or a lack thereof, based upon presence or increase or absence or decrease from normal cells of mRNA or protein from a gene shown to be up regulated or down regulated by a method of claim 13 in a mammal comprising:

- 30 (a) determining the level or status of mRNA in bone cells of said mammal; and/or  
(b) determining the level or status of corresponding protein in bone cells of said mammal; and



- (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining a presence or increase from normal cells, or an absence or decrease from normal cells, of mRNA or protein and thus risk.
16. A method for determining risk of developing osteoporosis or low or high bone density or other factors causing or contributing to osteoporosis or lower levels of osteoblasts and chondrocytes or other conditions involving mechanical stress or a lack thereof, based upon absence or decrease from normal cells of mRNA or protein from 608 comprising:
- (a) determining the level or status of mRNA in bone cells of said mammal; and/or
- (b) determining the level or status of corresponding protein in bone cells of said mammal; and
- (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining a presence or increase from normal cells of mRNA or protein and thus risk.
17. A method for determining risk of developing osteoporosis or low or high bone density or other factors causing or contributing to osteoporosis or of imbalance as to osteogenic and chondrogenic cells or other conditions involving mechanical stress or a lack thereof, based upon absence or decrease from normal cells or presence or increase from normal cells of mRNA or protein from 405 in a mammal comprising:
- (a) determining the level or status of mRNA in bone cells of said mammal; and/or

(b) determining the level or status of corresponding protein in bone cells of said mammal; and

- (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining an absence or decrease from normal cells of mRNA or protein and thus risk.

18. A method for determining risk of developing osteoporosis or low or high bone density or other factors causing or contributing to osteoporosis or of being susceptible to environmental factors or other than genetic factors of osteoporosis or of predisposition of bones towards susceptibility to environmental factors, or less lymphoid cells, or osteoporosis, or other conditions involving mechanical stress or a lack thereof, based upon presence or increase from normal cells or absence or decrease from normal cells of mRNA or protein from 274 in a mammal comprising:

- (a) determining the level or status of mRNA in bone cells of said mammal; and/or  
(b) determining the level or status of corresponding protein in bone cells of said mammal; and

- (c) comparing said level or status of mRNA and/or protein with the corresponding level in normal cells; wherein the term "level" denotes the amount of mRNA or protein produced; and, the term "status" includes that the gene, mRNA, protein or a transcription control element, including a promoter/enhancer sequence, may bear a mutation, deletion or any other modifications which would affect the overall activity of the gene when compared to the wild-type normal gene product, including post-translational modifications of the protein, and from the comparing determining an absence or decrease from normal cells of mRNA or protein and thus risk.

19. A method for testing a medicament for or gene therapy approach to osteoporosis or bone density or other factors causing or contributing to osteoporosis or symptoms thereof or other conditions involving mechanical stress or a lack thereof, based on absence or decrease from normal cells or presence or increase from normal cells of mRNA or protein comprising a method according to any one of claims 15-18

and additionally comprising: (a') administering the medicament or the gene therapy; and from the comparing determining an absence or decrease from normal cells or presence or increase from normal cells of the relevant mRNA or protein and thus efficacy of the medicament or the gene therapy.

5        20. A method for treating, preventing or controlling osteoporosis or other conditions involving mechanical stress or a lack thereof, comprising a method according to any one of claims 15-18 and further comprising administering a medicament or treatment for osteoporosis or a cause thereof or a symptom thereof.

21. A composition comprising a gene or portion thereof or a protein or  
10        portion thereof expressed by the gene or portion thereof or an antibody or portion thereof which binds to the protein or portion thereof, wherein the gene is identified by a method as claimed in claim 1 or 13.

22. An osteoporosis or mechanical stress or lack thereof model comprising  
15        bone cells which retain their characteristic thereof in culture with mechanical stress applied thereto or an absence of mechanical stress applied thereto.

23. An isolated nucleic acid molecule: encoding the herein identified  
protein 608 or a functional portion thereof or a polypeptide which is at least  
substantially homologous or identical thereto; or encoding the herein identified  
protein 405 or a functional portion thereof or a polypeptide which is at least  
20        substantially homologous or identical thereto; or encoding the herein identified  
protein 274 or a functional portion thereof or a polypeptide which is at least  
substantially homologous or identical thereto.

24. The isolated nucleic acid molecule of claim 23: encoding human  
protein 608 or a functional portion thereof; or encoding human protein 405 or a  
25        functional portion thereof; or encoding human protein 274 or a functional portion thereof.

25. The isolated nucleic acid molecule of claim 23: comprising SEQ ID  
NO: or a functional portion thereof; or comprising SEQ ID NO: or a functional  
portion thereof; or comprising SEQ ID NO: or a functional portion thereof; or  
30        comprising a nucleotide sequence as shown in the Figures or a functional portion thereof.

26. A vector comprising an isolated nucleic acid molecule of any one of  
claims 23-25.

27. A composition comprising the vector of claim 26.
28. A probe or primer which specifically hybridizes to an isolated nucleic acid molecule of any one of claims 23-25.
29. An expression product of the isolated nucleic acid molecule of any one of claims 23-25.
30. An isolated polypeptide: herein identified as protein 608 or a functional portion thereof or a polypeptide which is at least substantially homologous or identical thereto; or herein identified as protein 405 or a functional portion thereof or a polypeptide which is at least substantially homologous or identical thereto; or herein identified as protein 274 or a functional portion thereof or a polypeptide which is at least substantially homologous or identical thereto.
31. The isolated polypeptide of claim 30: which is human protein 608 or a functional portion thereof; or which is human protein 405 or a functional portion thereof; or which is human protein 274 or a functional portion thereof.
32. The isolated polypeptide of claim 28: comprising SEQ ID NO: or a functional portion thereof; or comprising SEQ ID NO: or a functional portion thereof; or comprising SEQ ID NO: or a functional portion thereof; or comprising a polypeptide sequence as shown in the Figures or a functional portion thereof.
33. A composition comprising the isolated polypeptide of any one of claims 30-32.
34. An antibody elicited by a polypeptide of any one of claims 30-32 or a functional portion thereof.
35. A composition comprising the antibody or functional portion thereof of claim 34.
36. A method for preventing, treating or controlling osteoporosis or bone density or other factors causing or contributing to osteoporosis or symptoms thereof or other conditions involving mechanical stress or a lack thereof, comprising administering a polypeptide as claimed in any one of claims 30-32.
37. A method for preventing, treating or controlling osteoporosis or bone density or other factors causing or contributing to osteoporosis or symptoms thereof or other conditions involving mechanical stress or a lack thereof, comprising administering a vector as claimed in claim 26.

38. A method for preventing, treating or controlling osteoporosis or bone density or other factors causing or contributing to osteoporosis or symptoms thereof or other conditions involving mechanical stress or a lack thereof, comprising administering a composition as claimed in claim 21.

- 5 39. A method for preventing, treating or controlling osteoporosis or bone density or other factors causing or contributing to osteoporosis or symptoms thereof or other conditions involving mechanical stress or a lack thereof, comprising administering a gene or functional portion thereof or a polypeptide comprising an expression product of the gene or functional portion of the polypeptide or an antibody  
10 to the polypeptide or a functional portion of the antibody, wherein the gene is identified in the model of claim 22.

40. The method of claim 39 wherein the gene identified in the model is CMF2-45 or CMF2-224.

- 15 41. A method for preparing a polypeptide comprising expressing the polypeptide from the vector of claim 26.

42. A method for preparing a polypeptide comprising expressing the polypeptide from a gene identified in a method as claimed in any one of claims 1 or 13.

- 20 43. A method for advancing research in or studies of bone development comprising a method as claimed in any one of claims 1 or 11.

44. A method for affecting a gene identified by any one of the methods of claim 1 or 11 comprising contacting cells containing the gene with a compound to which the gene responds.

- 25 45. The method of claim 44 wherein the compound comprises estrogen or a derivative or precursor thereof.

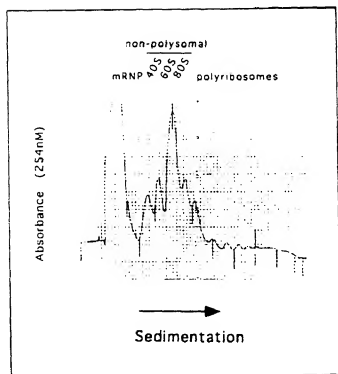


FIG. 1A

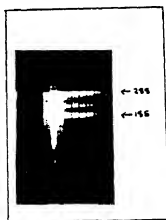
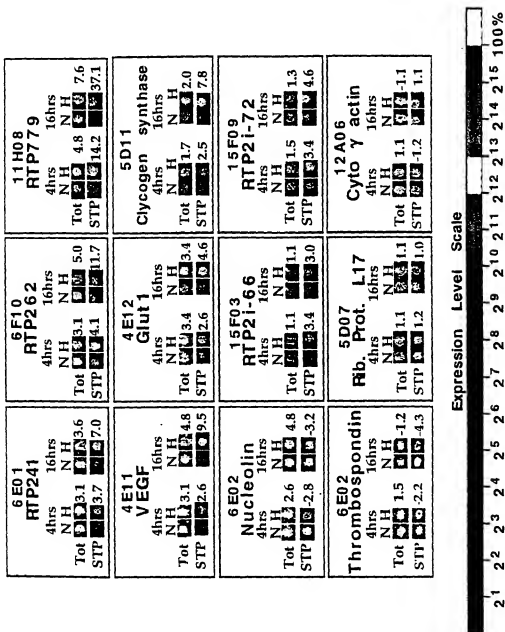


FIG. 1B

SUBSTITUTE SHEET (RULE 26)



**FIG. 2A**

[illegible]



## FIG. 2A-1

1.
 

>cal2-294 = a 2u globulin-related protein

TACTGTTTCTTTTCAGAGGTCTTCTGGAATAATACCATGGCAAACTGGTCGTAGTCAGTGTG  
TCGGCCCACTTGCACATCGTAGCTCTGTATCTGAGGGTAGCTGTGAATATTTCCCCAGGGTG  
AACTGGCCAGGCCCTGGAGCTTGGAAACGAATGTTCTGATCCAGTAGCGACGCCCTGGCCC  
CTCAGCAGGATGGAAAGTGACGTTGTAGCTATTGTCTTCTCTGTAGTCATAGATGGTGCTG  
T
2.
 

>call-42 = g-actin, cytoplasmic

ACATGGCTGGGGTATTGAAGTTTCAAACATTATCTGCGTCACTCTTCTCT  
CTGTAGCTTTGGGGTTCAAGGGGCGCTCGGTCAAGACACCGGGTGCTC  
CTCAGGGGCCACACGCAGCTCGTTGTAGAGGTGTGGTGCCAGATCTTCT  
CCATGTCTCCAGTTGGTGACAATGCCGTGCTCAATAGGGT
3.
 

>cal2\_54 = a-actin, cardiac

ACAGGTCTTTGCGGATATCGATGTCAACACTTCATGATGCTGTATATAGTAGTTTCGTGAA  
TGCCGGCAGACTCCATACCAATGAAGAAGGGCTGGAAAGAGTCTCAGGACACGCGGAAC  
GCTCGTTGCCAATGGTGATGACCTGGCCATCGGGAAGTTCGTAGCTCTTCTCCAGAGAGG  
AGGAAGATGACGCGGTGGCCATCTCATTCTCAAAATCCAGGGCGCATACACAGTTTCT  
CTTTGATGTCTCGGACAATTTTCAGTTTCAGCGGTGGTGACAAAGGAGT

## FIG. 2A-2

4.  
 >cal2\_224 = Alkaline phosphodiesterase  
 ACCNAGACCTGCCCCCTCGAATCAGACTCGTAATATACCTCAGACTTCTTTACCTTTA  
 ATTCCGAAGAGATTGTGAGAACCTCAGCTGCGAAAATCCGATCAACACTTCAAACCT  
 ACTTGACTCCGGATTGGCCGAACGACTGCNCTATGCCAAGAACGTGAGATTGACAAAG  
 TTCACTTAATGGTGATCGGCTATGGCTGGCTTACAGAACTAAGGAGATTCTCAACTGCG  
 AATCGGAANNACNGTTACACTATGAATTTAAANTTTGGNGCAATCTTTTGGCAC  
 ATGGACCCNGCTTTAANGAAGAACTGTTNTCCAAACCAATTTGAACAACATCGAAGTCTATA  
 NTCTACTGTGATCTTCTGCACNTCCN

5.  
 >cal1-582 = AM2 receptor/LDL receptor related ; LRP/a-2-  
 macroglobulin receptor  
 ACCCTTCTGCCTGCCCCCTGTTGAAGACCTGATGTCCTTTAAGCTCAGCGCCNATGCCCGT  
 TCTCAAGGACGAGTCTGTGGCATTGCCCTTGCATCCAGCTTGATGAGCCGTTGGC  
 GTGAGTCTGTCACTCCAGTAGATGAATCCTCAACACAGACAGAGAATGTCAT  
 GTTGTACTGGACAGAACCACTCAGGGTCTCAACCGCTCCAGATCGATACGTTGAT  
 CTTGTCCATCCGGGCATCACACAGTAGAGCTTGGCGCCTGATAGTCTACTGAGATGCC  
 ATTGGCCAGCTGATGCTGACATTTAACAGACCACTCTCTGTGCCCCGTCAGGCCAGA  
 CCGCTCAATCCGTGGGTAGTGAACCCCACTCAGTCCAGAACAG

# FIG. 2A-3

6.  $>cmf1-519 = \text{Aquaporin (AQP1)}; \text{Channel integral membrane protein}$   
 28;  
 ACCATGGGACCAACTGGCTGTTCCCTCTAGTCTCCCTTGACGTGCAGATGACGCTGTG  
 TGTATTATTAAGACACTGGGCTATTGACGGTCATGCTGAGGAAAGAAGCAGCTAGAC  
 ATGCAACAGAGCCCCAGACAGATGCCCCATGCCGGGCACACAGGGGTTGGATGCCCTAGT  
 TCATCATGATCAAGAGGGCCACTCTTGTCCTGGCTTATGAGCTGCTGGCTGATGAATTT  
 GACCAAGCCTGGACAACTGAAAGGGCTCACTATGTAAGCCCAAGAACATCTCAGAGTGCATTGAG  
 TGCAAGGACCTGATGCTGTGGCTCTGCTATAGCCCAAGAACATCTCAGAGTGCATTGAG  
 CTCAGGGCTTGACATTTAGCTCTCTGGGTTATTTCAATCCACCAAGCCAGAGCAGC  
 TCTACCACTGTGCCGTTAACCATGTCGTGAACGAGACCAATTCTTCAAGTGGCTTAGAA  
 GCAAGCAGAAATAATCAGAGGGCCATTGACCACTGGCATAAGT

6/183

SUBSTITUTE SHEET (RULE 26)

7.  
 $>cal1-138 = \text{Collagen type XII a-1}$   
 ACTTAAATGTGACAGATCGAAAACTTTACCAAGTTGATGGGACACCTTCTGTCTCAAT  
 GGTCACCTCAACCCGGCAGCGACCTCTTACAGGCTTAAAGTGAATCCTGCCGATGGAACCA  
 GGGGACAAAGAAATTAACAGTGAAGATCAGAGACCAATCACTGCTTCACTGGCCCTTCCC  
 CGGAGGCTGATGATGGCGTTACTGTTTTTGTGCAAGACACCAATCTCGAAGGGCCGGGTG  
 TCCCATCAAGAACAAGACAGACGACTGTGAACCAACAAGAGGCTCCACAGAACCCGCCACGC  
 CTTCACCTCCTCCCACTATTCACACCTGCCCGTGAATGATGCAAAAGGGGCCAAGGACAGATA  
 TTGTGTTCTGACGAGATGCCCTCTTGGAGTATGGAATGACAAATTTTAAACAAAGTTGTA  
 AATTATTTTAATACTGTGGGGCCTTGTATGAATCAAGCCCTGCTGGATTCAGGTTT  
 CTTTGTGACAGT

F1G. 2A-4

8. F10.2A + component C3

[illegible]

9. *subunit 1, mitochondria*

...cal2\_162 = Cytochrome oxidase subunit I, mitochondrial  
ACGATCCTGTTAGGCCCCCTACTGTGATTAAGAATACCTAAGCTCATATATG  
CGGGGGATCATTTGATATATCCCTCGCATAGTAGCAAGCACTGATTACTTTAAG  
CCTGTAGAGATTGGATATATAGTGGATGATGTAATTAACCTAAGGAAGCAATAGATATT  
AGGCTACTGTGATATGTGATGTGCTATACATTAACCTAAGGAAGCAATAGTACT  
ATGGCTCAACCATACCTATATATCCGAAGGTTCTTTTTCAGAGTAAAGTACT  
ACATGTGAATATATTCCAAACCTCGGAAGATTAAGATGT

## 10. FIG. 2A-5

>cal1-274 = DEST  
 ACCGTGTGCCACTTTAATGGAAGACAGCTTGATATATAACAAACGGCACT  
 TCAAGGTTATATACACACACGAGGATCACTCTAGGTAGTAGCCATC  
 AAACCTCCACTAAGCCAGACAGGATATGTAAATGTTGGATGGATGAT  
 TGGTGAAGATGTGGTTCGTGAGTCCCTCAGAATCTCCACGGCCTTCTGTGAG  
 T

## 11.

>cal2 144 = DEST  
 ACACACTGCCCACTCTCCCTCTGGCAGCAGCCCTGCCACTTCTTTACACTGCCAGCCAC  
 ATGTGGACATGGGCTACTCATGACTGTGGAAATTCCTTGGTGCCACCTTATTCCTCAGTG  
 GTGCTGTATGGCCCAATGGATGACATCAAGATATGCCAATGTGTTTGTGGGCCAGTT  
 GGTTCCTTGTGGGATCTGAGAGTGTGACTCCAGTCCAGGCTTTCAAGTTCTAATTC  
 ACTCTGTCTGAGCAAAACAGCTTCATGCCCGGATGAGTGAAGAACTATCCATGCTCAGTAA  
 TCTTCTGGCAGATCAGGGTTCTCTGTGTTTCATGAATGGAATGTGTCCTTCTCTCCA  
 CAACCAAAATTCATTCCTTCAAAAGAAAGACTTGCTGTGAGCCCACTCCAGAGGGTCTTTC  
 AAGGAGCCTTAGTCCCGAGAGTCCACTCTCCAGGTGAGCCCACTCCAGAGGGTCTTTC  
 CCACTGAATGGGTCTTCCAGCTGCTGGGAAAGTGGCTGTGNAGAACNCGCTCTTAAGACA  
 CAGTGGCTTCAGTTGCCAATGTTCAAGAGGAAGACCTCTTCTGGGGAATNANTGGGGA  
 GGGAAAGGATPAAAAATGACCTCCGTGTTTCTTGGTGTAACTTTCATTTGAATTTTAAATG  
 GCGANGCTCTCACACTGTTAACCACCTTAAAGANACCAATCACTTTCNNNGTGGCCATN  
 AAATGTTGAATGTCTTGGGCTCAGTTTCATTTCA

## FIG. 24-6

12.

>cell\_608 = DEST; homology to Neural cell adhesion molecule 1  
 ACCCTTGGTCCATATGACACATGTCATCAGCTTGGCTACATTTGTTGACACTGTCCCAT  
 CAGGCAAACTCCAGGATACCTTCAGGCACAGGGAGCCAGAGCCTTGAGTCACTTGA  
 AATCTTCCCATGGAGCAGCTGTCTTTAAATACTGTCTTCTGTCAATTTGGCAGGTG  
 TCAATCTCAGCGGACATGATCAGGACTAGGTCATCTCCATTTTGTTCGACAC  
 ATAGTAGTCACCAAGCCTTTTCCGTCACTGACCCAAACCAAGAGATCCATTTGGGT  
 AGACGTGGAATTCGGCTGCCCATTTCTGTGCACCTGTCGATGACAGCCTTGGATGGCAGCC  
 TCCAGATTATTTAGGCTTGGATCCCAAGTAGCTGAGCAGTTCAATGTAATTTCTAC  
 CCAATTCACCTCAGTCCCATGTCTGAGAGGCAATTCTATCTGGGGGTTGTCTCCCT  
 CTTCACAGTAGAAGACTACCTTCTCTGAGCCTGAGGAGCTGGTGGCAATGCACTC  
 ATAGTGCCTCTCACTGAAGAGCATGCTTCTTAATATACAGAGTTCATTTGGATACAG  
 AAAAATCTGGGAATGAGTCACTGCNATGCTTTAGTTCAGTCCCATCATAAAGAGACC  
 AGTGAACACTAGGGTGGGGAATTCCTTTGCACTGACAGGCAATTCAACTTCACCTA  
 AAAACCCCAAGATNGGCTTCTTTTGGATCCTAATGAACAGGGGAACNGTTGAA  
 TTAATCTTTTCTTTAAACCAACANTTGAATCTGNCANAAATGGGTGGCTGGCC

## FIG. 2A-7

13.  
 >cal2\_197 = DEST; some similarity to: erg3 (Ascobolus Immerusus;  
 Y10624)  
 ACCTGTCCCCACGAGCGCCGAGAAGTCGCAATTCACAGGCAATTCTTACAACTACAA  
 TGAATGGGATGATTCAACCCCCCGCATCGGAAGTGGTTGACTTCAAGCTGTTCTTCA  
 ATGGGCGCCCGGGGATGTGGCCCTGGACTCTTATCAACCTGTCTTTGCTGGCCAGAAGC  
 AGAGCTTTATGGCCATGTGACCACTCCATGATTTTGGTCAATGCTTACAGGCCATCT  
 ATGTGTAGACTTCTTGTGGAACGAAACCTGTAACA

14.  
 >cal1-405 = DEST=KIA0183  
 ACAATCAGTCCCTTCTGTGCTGTGCTCAGCCACTCTCAGCACTCTCAGCACTCAGGT  
 GCAACGGGGGGGAGGAGGTGTGTGATATCCATGTGTTCTCGTGGG  
 CATTGACAGGAGGACGGGATTTGGCTGAACAGTGCCCATGTGGCTCCCTT  
 TGCTTCAAGCTAGCTGGCTTCCCGAAGAGGCAGTGATGAACCTTCTAAC  
 TTTACTTGGTCTCCTGGGTCTCCCGTGCCAGTGGCTCTGCAATGGCTTCC  
 TGTCTTCTCCAGCCCAATCTGTTCCTCGCTGATGTGTTGGTTGCCCTC  
 CACTGCCCTCATCGTGTGTGATGAGGACGAGTAAACGAGCAGGGGAA  
 GTAGGGGAATCCAGTGATGGGGCTGCCGCTTGGGCCAAGTTGATGG  
 GCGAAGAAAGCTCATACAGCGGTGTGCTTCCCTCGTAGGAAATGT  
 TGCTGTAGCTGTTCTGCTCTGCAAGTTCTTCCCGTAGAGTCCAAGGCG  
 AAGCAGGCTCAGGGAAGCTGTATGGTATGAACCTGGCCCTGGGCTGCC  
 CTTGGGCTGGGACCTGTTGGAACCGCGGCTTGTGCTTGC

15. *FIG.24-8*

>cal14c8 = Endothelin converting enzyme  
 ACGGGACAAAGAACCTGTCTTCCCGCTGGAAGTTTGCCTGAGTGACACAGAAAACA  
 ACCCTGGGCTTTGGCCCTGGCCCCCATGTTGTGAAGCAACCTTTCGGAGACAGACAGA  
 ACATGACCAGCGAGATCATCTTGAGATCAAGAGGCATTCGAGAGAGCCCTGAGCACCC  
 TGAATGAGTGGATGAAGATACTCGGAGTCAGCCAAAGAGAGGGGAGCCCATCTACA  
 ACATGATAGGCTACCCCACTTCATCATGAGACCCCAAGAGATCAGTGAATTCAT  
 ACTACACAGCAGTCCCGATCTCTATTGAGAACGCCATGCGATTTTCACTTCAT  
 GAGGGTCACAGCCGACCAAGCTCAGGAAAGCCCCCAAGACGAGATTGTGTTCCAGCTGAA  
 CGCCCATGGTGAACCCCTACTACTCGGCCACCAAGACGAGATTGTGTTCCAGCTGAA  
 TCTTCAGAGGCGTATTATTATNCCCGCTCTTGGCCCAAGCCTTGAATTTGGTGTATCC  
 GGGCTTTGTTGGGCACNAAGTGAATCTGCTTCCCTTCNAAGTCAAGGCGNAGATTGAACA  
 GGATNGAANCTCCCGGCTGTTGGGAAAAACNCTCGGTGAAAGCTTCCANCCNCA  
 AGAATGCTTGGTTCAACCTTTCAAAACCTCCAACCCAAAGGGGGGCCCGGTTCC  
 ACCTTTGTCCCTTANTTGAAGGGTTAATTTCCAAACCTTGGGGTT

## 16.

>cal1-254 = filamin  
 ACTGAGCGGGCCAGGGTAGATGTCAGCATTTGATCAAGTGTGCCCCGTGAGTACTGGG  
 CCCCATGAGGCTGATATGTGACTTGTATCATCCGTAATGACAATGACACCTTCACTGT  
 GAATACACACCCCTGTGGGGTGGCAGCTAATACATCATGTCTTTTGTGACACAGC  
 CACACCCACAGCCCATCACACTCAATGGAAGCCTTTCATGATGCCAGTAAAGTGAA  
 TGCTGAGGCTCTGTCCCTAATGCACTGCTGTGAGCTTGGAAACCAACCATTTTAC  
 AGTCAATGCTAAACTGTCTGGGAAAGGCAAGCTGATGTTCAAGTTTCACGACTGGCTAA  
 CGAGATTTC



## FIG. 2A-9

17. >cal2 28 = Highly charged amino acid sequence (X59131)  
 ACTAGGATAGCGGGAGGCTAGGAGGAGAGACCACCAAGGCTCCCTGTCCACAG  
 TGCACAGTCAGTTGTTGAGATTCCACTGTCTGCAGGTCCTTGTCTCAATACCTGTT  
 AATAATATTTACATTGAATAATAGTCCCGAAGCCACTTTTGATATTCTAGGAGTCGG  
 GGAGTATTTCAAAGCTAGGGCCATCGGAGAGGGCCTTTGTTCAGGATTTTAGTCTTAG  
 GTCAATCAGGCCCTCCAACTTATCAGCCTTGAGCAACTCAAACAAGACTCCTATCCACCA  
 NCCTGGNCCCTCCCCCTCCCCCACCCTTTCTACCGTCTCAGGCATGGACACTGTGGTTTC  
 CTNAACCTGTGACTGATGCCCC
18. >cal2 204 = Kidney aminopeptidase M/kidney Zn-peptidase  
 aminopeptidase N  
 ACAGTAGTCCACTCCAGGCTCACAAGCCTTGCTAGTCGTTTCATTTCAGGCTGACCTGAT  
 CCTGTGACTGATGCAAGACGATTTCTCTCGGACAGCGGCACATTTTCATTGTGGCACA  
 GCTTGCTCAGCATTCCTCGTCTTCTCCGGATGATANAACCGGGTGCATGACTCCTCTA  
 GATTGTAGTAGGAGACCTTTGACCGACCCCGCTGGATAAGTCCACGTTAAGAAGT  
 GGTGGACTTTGAAGACAGGCACTCTCTTCGGAGTGTGAGATCTTTTCTAGGTAGATGA  
 TGAGGGTGTCTTGTGGAGAAGGCTTTGTCCATCTCATACTTGGAAATGATCTGTCTA  
 CTCCAGAGCTCAGNNGTNCAGGTCGTTGTGTGTGGAAATAAGCCANTCNTCATGGAGA  
 TGTCCAGGATGGACATAGTAGCATCCNCGTCTCCCAAGT

# FIG 2A-10

19.

>call-268 = Myosin heavy chain A, non-muscle  
 ACTCTCTCGTCCAGNATGAACATGGTGTGGTTGACAGCTGCTGCAGC  
 TTCTCGTNGTWTACTNGATNTCAAAGTTGCTCGAATGAGTTACATCAA  
 AGATCTCGAAGCCAGCGGATGTCCANGATCNCANNATGANGCNCCCTG  
 CCTCTGGTCTTGTCCNNTAACTTTGTTATTNGAACACCAGCCAGGGGA  
 ATATCCGCNCGTAGGTACCTTNGNCAGAGCCTCCATGGTAAATCAGCC  
 TGCNCTTTCTGTGTGGCCNTCTGCNCATNCTCGTGCCTNCTNATGNC  
 CGGGTCAAGGATNCCCTGTGTTGACCTCGGTGTCNCTNATCNCNCCAC  
 GCAAGTNGGACANCTNTGAACCTCCTGTGTATCTGGCATGGACGACTT  
 GTTCAGNTGCTCCCTCCCTNCTNCTGANNGACCCANGTNCCTCTGTG  
 TNATGTACCCCTTGANATGATCTCTCAG

20.

>call14d6t7 = Novel protein AHNAK (nucleoprotein) (M80899)  
 ACATTTTTCAAACATGGTTTCTTGACTTTGGGAGATTTTACTCAGCTTGAGCGCAG  
 CATCAGGCCCTGCAACGCTGACATCCACATCAGGAATTTGAGGCTGGCATCAATCTCAC  
 CTCAGNAACATTCAAATCAATCCCTGTTCTTGCTCTTGACTTTAGGGCCACTCATGT  
 GAATTTAGGCATTTTAACTTGCTTTTCTTGCCCTTGCCACCAATGTTTATCTCAGGAG  
 TCTCAACACTATTTCTGTTTCAGGGACAGTCACACCAACTGTGGGGTCTTAATGTCAG  
 CTTTGGGGCTTTTGCACCAATCCAACTTGGTTTAACTTCGAACTTTAAACAT  
 CTGGCCCTTCAATATTAATATCTGGACTCTCCATGTGGACATCTAAGCTTGGAGTTTCTA  
 AATCAACTTTAGGCCCCCTGAGGTCAACTTCACCTCTCTAATAGGCCCCAGAAATTC  
 CAACATTTGGTGTCTTGAGGTGCACATTAGCATCAGGCCCACTTATTTTAGGAACAGGT

## FIG. 2A - 11

21. >call-191 = Secretory protein containing thrombospondin motifs

ACGAGACCTTGTGCAGACCTTCTTGCCACAGTTGGCAGGTGGGGATTG  
 GTCACCATGTTCTAAAACTTGTGGGAAGGTTATAAGAAGAACCTTGA  
 AATGCTGTCCCATGATGGGGTGTGTATCAATGAGAGCTGTGATCCT  
 TTGATATAACCAAGCATTACATTTGACTTTTGCTACTGACACAGTGCAG  
 TTAACAGGTTTTGAGGACAAGGTAGCGTGGGAGGCTGATACACTGAAAG  
 CAAGAGTGCTGGAGGGATCCCTGTGAATCAAACCACTAAACAGTGAGGTGT  
 GGCAATGAGGTGTGTATAGGGGATACATTAGCAAAAGAGGTAGATCCGT

22.

>call-562 = Stearoyl Co A desaturase

ACTGTGCAGCGGGCTGGCTGCAGAACTTAAGTCATAAACACTCTTTGCTGAGGTCTG  
 AAGTCTCTGCACGTTCTCATCCCTGGGAACGTGACCCAGCATCCGACGCCAAGATGCC  
 GGCTCACATACTGCAAGAGATCTCTGGCTTACTCGGCCACCAACACAAATCAGAGCGCC  
 ACCTTCTGGGGGACAGAGATGGAGGAGAGAAAGTTTGAAGGAAATCCTCACCACTGGGG  
 AGCAGATGTTCTGCCCTGAAATTAAGATGACCTATACGACCCCGCTACCGGATGAGGA  
 GGGGCCCCCGCCCAAGCTGGAGT

# FIG. 2A-12

23. >cal2\_76 = Tenascin  
 CATTTGGAGCTCCACCGGCGGTGGCGGCTCTAGATAAGTGGATCCCCGGGCTGCAGGA  
 ATTGGATATCTCATGGGCCCTTCCAGTTGGAAACCATTTAAACCCCTGAATGGTGAATA  
 ATGTTCCCATATCTGCCCATCAAGGTTGAANCCATGANAGTTCTTATACCAAGAACTCC  
 TTTGTNGGACACGGGCACAGTTGGTGAATGGCTGAATNGTGTCTTGTCTATAGGTNGANAA  
 NGATCTACCAATTTGATAGTTTCATGGATCACTACCTACCGTTCCACTGTATCCTTCTACCTT  
 CAGCTGTAGCGACTCTTGGCATCTCCAACTCACTGAACCTGTCTATACACAGCATAAAGCTGA  
 CTCCCATGGTCTTGTAGGTCCACCCGGANCTCATATTGCCCTTGGGCTGTGATTTTCT  
 CAGTTATCCAGTCCAAAGCAGAAATTCCTCTCTGCGGTCTCCAAACCCAATGGCATTAGC  
 CTTCCAGTTTCCAAAGAAANTCCTCACGTCCATTTTGGCTCTCAAGAAACGANTCATC  
 CACTCCATCC
24. >call-609 = Thrombospondin 1  
 ACCACTGAAGTCCACGGCATTAATACTCATCATACCTACAGCAAGTCCAGGTCACAGTT  
 TACAGTCTGGACAAGTCTTTGCCCTGATGGCGGACAAACCACTAGGTCATTTTGGGA  
 GGTTCCTTTGGGATCTAGAGGAATCATCTGGAATCGGCGGAAATCGGTTTCACTGATGTC  
 AACATCTCAGGACAGATGTCTCAATGTCTGGCAATTTGTCTGGTCAAAAGTCGTCTTT  
 GCAGGCATCGCTCGGCCATCACCATCAGAGTCTTCTGGTCAGGATTTGGGCACACGCT  
 GCAGTTGTCTCTGTCTATCAGGATGCCGTGTTGTCTATCGTCATGGTCAACAGGCGTCTCC  
 TTTTACCATCTTATCGTGTGCGGCTGGTTGGCATTTGGGCACATAGGACAGTTGTGTCCAC  
 GTTGTCTGTATGGCCATCTCATCGATGTCTCTGATTTGTGTC

## FIG. 2A-13

25.

&gt;call-700 DEST

ACTCTCCCAAGGTGTTGAGGTGGGAGGAAAGGGGCTTCTGCACCTTCGATGATGGCCTCG  
 ATACTCTGACTGCGGGTTCTGCAGCTCTGTTTGGTAGCACCTGCTGGCAGGAAACTTCTCTG  
 CCTCTGGATGAAAGGAGGGTTACTCATCATCATCTGGCAGCTGCAGGAANTCGGGGGCC  
 TGATAAACAAATTTTCAACCCCTTTCTGCTAGTCCGCTCTGCCAAAAAACTCAAAGGACCC  
 AAGCAGGGCAGCTGTTCCACAGCGAGCCACTGTCTAGCTGAGT

26.

&gt;call-248 Connective tissue growth factor

ACTACAGGAAGATGTATGGAGACATGGCGTTAAGCCAGGGAGTAGGGACACGAACTCAT  
 TTAGACTATTAACCTTGAAGTGAATTAATCTGCTCATTTCTTCTGTAAATAAAAGGAT  
 TACAGTAGCACATTAATTAATCTGGGTTCTTAACCTGCTGGGAGAAACACCCACC  
 GAAGTAGAACCCTGTTGTCATTTGTCATGCAATAGCCCTGTCAATCTCAGACACTGGTTTC  
 GAGACAGTTTAGACTTGACAGTTGTTCACTAGCCACAGTGACAGAACGCACACTAAGGT  
 GAGCCTCCTGGCAGAGTGGACATGCCANCGAGAAAGACAGGT

27.

&gt;cal2-45 SARP1

ACTACAGCTATTCAAACACTACCCTTAAATGAGCTGTTTTAAACCTTCATAAAAAACAGT  
 CATGATTTTATTAGTTGTGAATATTTCTACAAAGATCTGGATGGACTTTTCGCTTTATGTG  
 AAAATAGCTATCTGTTCCCGGGGCCCTCTAGGGCCTTATAACTCAGGACACGTTAGGATA  
 CACACATGCAAAAGGCGAGGSCAAGCTGCTCCAAAGCTGAGACTGTAGCTCTCCCGACAGT

## FIG. 2A-14

28. >call3c4 Syntrophin  
 ACACAGAAGCCATTCCGAGGCTGGACCTGACCCCTGTGGGATCCACCCCATCTCCAGTC  
 AACAGCAGAGCTGCCACACCTCCTGGAGGCCCTTAAGCCACCCCTGAGATCTCTGCAGATG  
 GAGTCTTCTGGGTCAAGTGCACAAAGGAGGAGGAGGCTTTCTGTTCACTCTCCCTTGG  
 GCCCGAGGGCTGTCTGGGCTCTCTCCTCTCTCTCCTTCCCTTGTGATAATAACAGC  
 CTGTGAGCACATAGACAAAAAATAAAAAAAGCTTGT
29. >call-404 erg25 methyl sterol  
 ACTGGACATCCGTCCCAAGATTCTGTCCACCACCGTGAAGTGGAGGCTAATTCCCAA  
 TGAATTTCATGTGGTGAATCTGTATGCCGAGCCCCGGTATAAAAAGGGATGTAGTTCA  
 ACGGTTGAGAGGGATATCAAAACCACTGTGCACATCGATGGTTCTAGCAACCGCATGG  
 TCACCCATGCCACAGAAGATTACATGATCACACAAAAGCAATTCGAATGAAAAATGGA  
 CCAGTCCCAAGATTAGGTTTCCAAAGGATGTGCATATTTCTGCTTCGATTCCAATGGA  
 GCCTGGAAACTCTGTGATGGACTTTATGAATGATTTATAAATCCTCTTGTGGTGAAGGAG  
 TCTGTGCAGGAATAGTGCCAGGTATCCTCGATGACTGCACAGCCCAACACCTTGCCAA  
 AGTGAATACCATCTTGGCAATCTTTCCCAATCAATGAAGGAATATTGAAAAACTCTGTAA  
 ATAGTATGTTCCCAAAATCAGAACCACTGGATAAAAAATGATTAACAGAAATTCCTTT  
 CAAACACTTCCACTGGCCCTTCAAAGGTTCTGGTTTATCCTTTTGAATCTTGT

*FIG. 2A-15*

30.

&gt;cal2-176 DEST

ACCATAAGCAGTTTCCATGGGGTCAGAGAATGGCTGGAGAGTTGTGGAAACCCAGCAGGAG  
GTATTGTTTGGAGACAGGACCTCACTGTGTAGCCCTGGCCCTCAAACTCAGAGATCCACC  
TGCTCTGCCCTCCTGAGTGCTTGGATTAAAGGCATGCACCCACCCGCCCTGACAGGG  
AGGAAATTCTGATGGAGGAGCTATTGACAGGTAGCAGACAGACCCGGAAGACGATCATG  
GGCCATCTGTGGGTCAGTGGGGCTGCTGTATGTCTTTGCTTTACAGGGGAAGGTGA  
TGTATGCATCGTAGCCAAACAGCAAGGTAGCCANTAAAGCCCAAGT

31.

&gt;OSTEONECTIN

ACTTGTCAATTGCTTAGGTACAGGTCTCAAAGAAAGCGAGTGGTGCATGTTCATGGGAA  
TGAGGGGAGCGGCAGTGGGGCCAGCTCCGTGTGGGACAGGT

## FIG. 3A

## Protein Sequence of 608

Length of 608 13.5.99: 8883 bp; Listed from: 2 to: 8883;  
Translated from: 575 to: 8365 (Entire region);  
Genetic Code used: Universal;

## Frame 2

GAG AGA CGA CAG AAG GTT ACG GCT GCG AGA AGA CGA CAG AAG GGT CCA GAA AAA  
10 19 28 37 46 55

GGA AAG TGC TGG AGG GGA GTG GGG ACA AAA GCA GCG ACC AAG TGA ATG TCA CTT CAG TGA  
64 73 82 91 100 109

CTG AGG CCA GGC AAA ACG CSC GGG AAG GAT TTT GTG TAG CTT GGG ACC CTT TCA TAG ACA  
121 133 142 151 160 169

CTG ATG ACA CGT TTA CCG AAA ATA GAA ATT TGA GGA AAG CCG CCT GGG CCT TCG GAA AGG  
184 193 202 211 220 229



## FIG. 3B

AGT GAT TGA TTA GTA CTT GCA AGT TTA GGT GAC TTT AAG GAG AAC TAA CTA ATG TAT ACT  
244 253 262 271 280 289

ATT GAG GGA GGA GGA AGA GCA TTA CAG AGT TTC CAG CAG CAG GAA AGC TTT GGT TAA  
304 313 322 331 340 349

TTT GGA AAT GGA TGA TAG CAT TAA AAT AAC AGA AGC GGC TCC AGG TCT CTG AAG CTT CAG  
364 373 382 391 400 409

TCC CCC AGC TGA AAG CCA GAA AAG ACT AAG CCC ACT AAG CCT TTT GAT CCC TTT GGA AGC  
424 433 442 451 460 469

AAA GAA CTT TCC TTC CCT GGG GTG AAG ACT CTC CTC AGA AGA TTT CCT GTC TCT GCC TAT  
484 493 502 511 520 529

## FIG. 3C

GTT ACA AGA GGA ATC AAA ACC AAG ACA GAA GAG CTC AGG ATG CAG GTG AGA GGC AGG GAA	544	553	562	571	580	589
V S G L L I S L T A V C L V T P G S R	604	613	622	631	640	649
GTC AGC GGC TTG TTG ATC TCC CTC ACT GCT GTC TGC CTG GTG GTC ACC CCT GGG AGC AGG						
A C P R R C A C Y V P T E V H C T F R Y	664	673	682	691	700	709
GCC TGT CCT CGC CGC TGT GCC TGC TAT GTG CCC ACA GAG GTG CAC TGT ACA TTT CGG TAC						
L T S I P D G I P A N V E R I N L G Y N	724	733	742	751	760	769
CTG ACC TCC ATC CCA GAT GGC ATC CCG GCC AAT GTG GAA CGA ATA AAT TTA GGA TAT AAC						
S L T R L T E N D F D G L S K L E L L M	784	793	802	811	820	829
AGC CTT ACT AGA TTG ACA GAA AAC GAC TTT GAT GGC CTG AGC AAA CTG GAG TTA CTC ATG						

22/183

## FIG. 3D

L H S N G I H R V S D K T F S G L Q S L  
 CTG CAC AGT AAT GGC ATT CAC AGA GTC AGT GAC AAG ACC TTC TCG GGC TTG CAG TCC TTG  
 844 853 862 871 880 889  
  
 Q V L K M S Y N K V Q I I R K D T F Y G  
 CAG GTC TTA AAA ATG AGC TAT AAC AAA GTC CAA ATC ATT CGG AAG GAT ACT TTC TAC GGA  
 904 913 922 931 940 949  
  
 L G S L V R L H L D H N N I E F I N P E  
 CTC GGG AGC TTG CTC CGG TTG CAC CTG GAT CAC AAC AAC ATT GAA TTC ATC AAC CCT GAG  
 964 973 982 991 1000 1009  
  
 A F Y G L T S L R L V H L E G N R L T K  
 GCC TTT TAT GGA CTT ACC TCG CTC CGC TTG GTA CAT TTA GAA GGA AAC CGG CTC ACA AAG  
 1024 1033 1042 1051 1060 1069  
  
 L H P D T F V S L S Y L Q I F K T S F I  
 CTC CAT CCA GAC ACA TTT GTC TCA TTA AGC TAT CTC CAG ATA TTT AAA ACC TCT TTC ATT  
 1084 1093 1102 1111 1120 1129

## FIG. 3E

K Y L F L S D N F L T S L P K E M V S Y  
 AAG TAC CTG TTC TTG TCT GAT AAC TTG CTG ACC TCC CTC CCA AAA GAA ATG GTC TCC TAC  
 1144 1153 1162 1171 1180 1189  
  
 M P N L E S L Y L H G N P W T C D C H L  
 ATG CCA AAC CTA GAA AGC CTG TAT TTG CAT GGA AAC CCA TGG ACC TGT GAC TGC CAT TTA  
 1204 1213 1222 1231 1240 1249  
  
 K W L S E W M Q G N P D I I K C K K D R  
 AAG TGG TTG TCT GAG TGG ATG CAG GGA AAC CCA GAT ATA ATA TGC AAG AAA GAC AGA  
 1264 1273 1282 1291 1300 1309  
  
 S S S S Q C P L C M N P R I S K G R  
 AGC TCT TCC AGT CCT CAG CAA TGT CCC CTT TGC ATG AAC CCC AGG ATC TCT AAA GGC AGA  
 1324 1333 1342 1351 1360 1369  
  
 P F A M V P S G A F L C T K P T I D P S  
 CCC TTT GCT ATG GTA CCA TCT GGA GCT TTC CTA TGT ACA AAG CCA ACC ATT GAT CCA TCA  
 1384 1393 1402 1411 1420 1429

## FIG. 3F

L K S K S L V T Q E D N G S A S T S P 2  
 CTG AAG TCA AAG AGC CTG GTT ACT CAG GAG GAC AAT GGA TCT GCC TCC ACC TCA CCT CAA  
 1444 1453 1462 1471 1480 1489  
  
 D F I E P G S L S L N M T X X S G N K  
 GAT TTC ATA GAA CCC TTT GGC TCC TTG TCT TTG AAC ATG ACA KAS MTK TCT GGA AAT AAG  
 1504 1513 1522 1531 1540 1549  
  
 A D M V C S I Q K P S S R T S P T A F T E  
 GCC GAC ATG GTC TGT AGT ATC CAA AAG CCA TCA AGG ACA TCA CCA ACT GCA TTC ACT GAA  
 1564 1573 1582 1591 1600 1609  
  
 E N D Y I M L N A S F S T N L V C S V D  
 GAA AAT GAC TAC ATC ATG CTA AAT CCG TCA TTT TCC ACA AAT CTT GTG TGC AGT GTA GAT  
 1624 1633 1642 1651 1660 1669  
  
 Y N H I Q P V W Q L L A L Y S D S P L I  
 TAT AAT CAC ATC CAG CCA CTG TGG CAA CTT CTG GGT TTA TAC AGT GAC TCT CCT CTG ATA  
 1684 1693 1702 1711 1720 1729

## FIG. 3G

L E R K P Q L T E T P S L S S R Y K Q V  
 CTA GAA AGG AAG CCC GAG CTT ACC GAG ACT CCT TCA CTG TCT TCT AGA TAT AAA CAG GTG  
 1744 1753 1762 1771 1780 1789  
  
 A L R P E D I F T S I E A D V R A D P F  
 GCT CTT AGG CCT GAA GAC ATT TTT ACC AGC ATA GAG GCT GAT GTC AGA GCA GAC CCT TTT  
 1804 1813 1822 1831 1840 1849  
  
 W F Q Q E K I V L Q L N R T A T T L S T  
 TGG TTC CAA CAA GAA AAA ATT GTC TTG CAG CTG AAC AGA ACT GCC ACC ACA CTT AGC ACA  
 1864 1873 1882 1891 1900 1909  
  
 L Q I Q F S T D A Q I A L P R A E M R A  
 TTA CAG ATC CAG TTT TCC ACT GAT GCT CAA ATC GCT TTA CCA AGG GCG GAG ATG AGA GCG  
 1924 1933 1942 1951 1960 1969  
  
 E R L K W T M I L M N N P K L E R T V  
 GAG AGA CTC AAA TGG ACC ATG ATC CTG ATG ATG AAC AAT CCC AAA CTG GAA CGC ACT GTC  
 1984 1993 2002 2011 2020 2029

## FIG. 3H

L V G G T I A L S C P G K G D P S P H L	
CTG GTT GGC GGC ACT ATT GCC CTG AGC TGT CCA GGC AAA GGC GAC CCT TCA CCT CAC TTG	2089
	2044
	2053
	2062
	2071
	2080
E W L L A D G S K V R A P Y V S E D G R	
GAA TGG CTT CTA GCT GAT GGG AGT AAA GTG AGA GCC CCT TAC GTT AGC GAG GAT GGG CGA	2149
	2104
	2113
	2122
	2131
I L I D K N G K L E L Q M A D S F D A G	
ATC CTA ATA GAC AAA RAT GGG AAG TTG GAA CTG CAG ATG GCT GAC AGC TTT GAT CCA GGT	2209
	2164
	2173
	2182
	2191
	2200
L Y H C I S T N D A D A D V L T Y R I T	
CTT TAC CAC TGC ATA AGC ACC NAT CAT CCA GAT GCG GAT GTT CTC ACA TAC AGG ATA ACT	2269
	2234
	2242
	2251
	2260
V V E P Y G E S T H D S G V Q H T V V T	
CTG GTA GAG CCC TAT GGA GAA AGC ACA CAT GAC AGT GGA GTC CAG CAC ACA GTG GTT ACG	2329
	2284
	2293
	2302
	2311

## FIG. 3I

G E T L D L P C L S T G V P D A S I S W  
 GGT GAG ACG CTC GAC CTT CCA TGC CTT TCC AGG GGT GTT CCA GAT GCT TCT ATT AGC TGG  
 2344 2353 2362 2371 2380 2389  
  
 I L P G N T V F S Q P S R D R Q I L N N  
 ATT CTT CCA GGG AAC ACT GTG TTC TCT CAG CCA TCA AGA GAC AGG CAA ATT CTT AAC ATT  
 2404 2413 2422 2431 2440 2449  
  
 G T L R I L Q V T P K D Q G H Y Q C V A  
 GGG ACC TTA AGA ATA TTA CAG GTT ACG CCA AAA GAT CAA GGT CAT TAC CAA TGT GTG GCT  
 2464 2473 2482 2491 2500 2509  
  
 A N P S G A D F S S F K V S V Q K K G Q  
 GCC AAC CCA TCA GGG GCC GAC TTT TCC AGT TTT AAA GTT TCA GTT CAA AAG AAA GGC CAA  
 2524 2533 2542 2551 2560 2569  
  
 R M V E H D R E A G G S G L G E P N S S  
 AGG ATG GTT GAG CAT GAC AGG GAG GCA GGT GGA TCT GGA CTT GGA GAA CCC AAC TCC AGT  
 2584 2593 2602 2611 2620 2629  
  
 V S L K Q P A S L K L S A S A L T G S E  
 GTT TCC CTT AAG CAG CCA GCA TCT TTG AAA CTC TCT GCA TCA GCT TTG ACA GGG TCA GAG  
 2644 2653 2662 2671 2680 2689



## FIG. 3J

A G K Q V S G V H R K N K H R D L I H R  
 GCT CGA AAA CAA GTC TCC GGT GTA CAT AGG AAG AAC AAA CAT AGA GAC TTA ATA CAT CGG  
 2704 2713 2722 2731 2740 2749

R R G D S T L R R F R E H R R Q L P L S  
 CGG CGT GGG GAT TCC AGC CTC CGG CGA TTC AGG GAG CAT AGG AGG CAG CTC CCT CTC TCT  
 2764 2773 2782 2791 2800 2809

A R R I D P Q R W A A L L E K A K K N S  
 GCT CGG AGA ATT GAC CCG CAA CGC TGG GCA GCA CTT CTA GAA AAA GCC AAA AAG AAT TCT  
 2824 2833 2842 2851 2860 2869

V P K K Q E N T T V K P V P L A V P L V  
 GTG CCA AAA AAG CAA GAA AAT ACC ACA GTA AAG CCA GTG CCA CTG GCT GTT CCC CTC GTG  
 2884 2893 2902 2911 2920 2929

E L T D E E K D A S G M I P P D E E F M  
 GAA CTC ACT GAC GAG GAA AAG GAT GCC TCT GGC ATG ATT CCT CCA GAT GAA GAA TTC ATG  
 2944 2953 2962 2971 2980 2989

V L K T K A S G V P G R S P T A D S G P  
 GTT CTG AAA ACT AAG GCT TCT GGT GTC CCA GGA AGG TCA CCA ACT GCT GAC TCT GGA CCA  
 3004 3013 3022 3031 3040 3049

FIG. 3K

V N H G F M T S I A S G T E V S T V N P  
 GTA AAT CAT GGT TTT ATG ACG AGT ATA GCT TCT GGC ACA GAA GTC TCA ACT GTG AAT CCA  
 3064 3073 3082 3091 3100 3109

Q T L Q S E H L P D F K L F S V T N G T  
 CAA ACA CTA CAA TCT GAG CAC CTT CCT GAT TTC NAA TTA TTT AGT GTA ACA AAC GGT ACA  
 3124 3133 3142 3151 3160 3169

A V T K S M N P S I A S K I E D T T N Q  
 GCT GTG ACA AAG AGT ATG AAC CCA TCC ATA GCA AGC NAA ATA GAA GAT ACA ACC AAC CAA  
 3184 3193 3202 3211 3220 3229

N P I I I F P S V A E I R D S A Q A G R  
 AAC CCA ATC ATT ATC TTT CCA TCA GTA GGT GAA ATT CGA GAT TCT TCT GCT CAG GCA GGA AGA  
 3244 3253 3262 3271 3280 3289

A S S Q S A H P V T G G N M A T Y G H T  
 GCA TCT TCC CAA AGT GCA CAC CCT GTA ACA GGG GGA AAC ATG GCT ACC TAT GGC CAT ACC  
 3304 3313 3322 3331 3340 3349

N T Y S S F T S K A S T V L Q P I N P T  
 AAC ACA TAT AGT AGC TTT ACC AGC AAA GCC AGT ACA GTC TTG CAG CCA ATA AAT CCA ACA  
 3364 3373 3382 3391 3400 3409

## FIG. 3L

E S Y G P Q I P I T G V S R P S S S D I  
 GAA AGT TAT GGA CCT CAG ATA CCT ATT ACA GGA GTC AGC AGA CCT AGC AGT ACT GAC ATC  
 3424 3433 3442 3451 3460 3469  
  
 S S H T T A D P S F S S H P S G S H T T  
 TCT TCT CAC ACT ACT GCA GAC CCT AGC TTC TCC AGT CAC CCT TCA GGT TCA CAC ACC ACT  
 3484 3493 3502 3511 3520 3529  
  
 A S S L F H I P R N N T G N F P L S R  
 GGC TCG TCT TTA TTT CAC ATT CCT AGA AAC AAC ATT ACA GGT AAC TTC CCC TTG TCC AGG  
 3544 3553 3562 3571 3580 3589  
  
 H L G R E R T I W S R G R V K N P H R T  
 CAC TTG GGA ATA GAG AGG ACA ATT TGG AGC AGA GGG AGA GTT AAA AAC CCA CAT AGA ACC  
 3604 3613 3622 3631 3640 3649  
  
 P V L R R H R H R T V R P A I K G P A N  
 CCA GTT CTC CGA CGG CAT AGA CAC AGG ACT GTG AGG CCA GCA ATC AAG GGA CCT GCT AAC  
 3664 3673 3682 3691 3700 3709  
  
 K N V S Q V P A T E Y P G M C H T C P S  
 AAA ATT GTG AGC CAA GTT CCA GCC ACA GAG TAC CCT GGG ATG TGC CAC ACA TGT CCT TCC  
 3724 3733 3742 3751 3760 3769

## FIG. 3M

A. E G L T V A T A A L S V P S S H S A  
 GCA GAG GGG CTC ACA GTG GCT ACT GCA GCA CTG TCA TCC CAC AGT GCC  
 3784 3793 3802 3811 3820 3829

L P K T N N V G V I A E S T T V V K K  
 CTC CCC AAA ACT AAT AAT GTT GGG GTC ATA GCA GAA GAG TCT ACC ACT GTG GTC AAG AAA  
 3844 3853 3862 3871 3880 3889

P L L L F K D K Q N V D I E I I T T T  
 CCA CTG TTA CTA TTT AAG GAC AAA CAA AAT GTA GAT ATT GAG ATA ATA ACA ACC ACT ACA  
 3904 3913 3922 3931 3940 3949

K Y S G G E S N H V I P T E A S M T S A  
 AAA TAT TCC GGA GGG GAA AGT AAC CAC CAC GTG ATT CCT ACG GAA GCA AGC ATG ACT TCT GCT  
 3964 3973 3982 3991 4000 4009

P T S V S L G K S P V D N S G H L S M P  
 CCA ACA TCT GTA TCC CTG GGG AAA TCT CCT GTA GAC AAT AGT GGT CAC CTG AGC ATG CCT  
 4024 4033 4042 4051 4060 4069

G T I Q T G K D S V E T T P L P S P L S  
 GGG ACC ATC CAA ACT GGG AAA GAT TCA GTG GAA ACA ACA CCA CTT CCC AGC CCC CTC AGC  
 4084 4093 4102 4111 4120 4129

FIG. 3N

T P S I P T S T K F S K R K T P L H Q I  
 ACA CCC TCA ATA CCA ACA AGC ACA AAA TTC TCA AAG AGG AAA ACT CCC TTG CAC CAG ATC  
 4144 4153 4162 4171 4180 4189  
  
 F V N N Q K K E G M L K N P Y Q F G L Q  
 TTT GTA AAT AAC CAG AAG AAG GAG GGG ATG TTA AAG AAT CCA TAT CAA TTC GGT TTA CAA  
 4204 4213 4222 4231 4240 4249  
  
 K N P A A K L P K I A P L L P T G Q S S  
 AAG AAC CCA GCC GCA AAG CTT CCC AAA ATA GGT CCT CTT TTA CCC ACA GGT CAG AGT TCC  
 4264 4273 4282 4291 4300 4309  
  
 P S D S T T L L T S P P P A L S T T M A  
 CCC TCA GAT TCT ACA ACT CTC TTG ACA AGT CCG CCA CCA GCT CTG TCT ACA ACA ATG GCT  
 4324 4333 4342 4351 4360 4369  
  
 A T Q N K G T E V V S G A R S L S A G K  
 GCC ACT CAG AAC AAG GGC ACT GAA GTA GTA TCA GGT GCC AGA AGT CTC TCA GCA GGG AAG  
 4384 4393 4402 4411 4420 4429  
  
 K Q P F T N S S P V L P S T I S K R S N  
 AAG CAG CCC TTC ACC AAC TCC TCT CCA GTG CTT CCT AGC ACC ATA AGC AAG AGA TCT AAT  
 4444 4453 4462 4471 4480 4489

## FIG. 30

T L N F L S T E T P T V T S P T A T A S  
 ACA TTA AAC TTC TTG TCA ACG GAA ACC CCC ACA GTG ACA AGT CCT ACT GCT ACT GCA TCT  
 4504 4513 4522 4531 4540 4549  
  
 V I M S E T Q R T R S K E A K D Q I K G  
 GTC ATT ATG TCT GAA ACC CAA CGA ACA AGA TCC AAA GAA GCA AAA GAC CAA ATA AAG GGG  
 4564 4573 4582 4591 4600 4609  
  
 P R K N R N A N A N T P R Q V S G Y S A  
 CCT CGG AAG AAC AGA AAC AAC GCA AAC ACC ACC CCC AGG CAG GTT TCT GGC TAT AGT GCA  
 4624 4633 4642 4651 4660 4669  
  
 Y S A L T T A D T P L A F S H S P R Q D  
 TAC TCA GCT CTA ACA ACA GCT GAT ACC CCC TTG GCT TTC AGT CMT TCC CCA CGA CAA GAT  
 4684 4693 4702 4711 4720 4729  
  
 D G G N V S A A V A Y H S T T S L L A I T  
 GAT GGT GGA AAT GTA AGT GCA GTT GCT TAT CAC TCA ACA ACC TCT CTT CTG GCC ATA ACT  
 4744 4753 4762 4771 4780 4789  
  
 E L F E K Y T Q T L G N T T A L E T T L  
 GAA CTG TTT GAG AAG TAC ACC CAG ACT TTG GGA AAT ACA ACA GCT TTG GAA ACA ACG TTG  
 4813 4822 4831 4840 4849

FIG. 3P

L S K S Q E S T T V K R A S D T P P L  
 TTG AGC AAA TCA CAG GAG AGT ACC ACA GTG AAA AGA GCC TCA GAC ACA CCA CCA CTC  
 4864 4873 4882 4891 4900 4909  
  
 L S S G A P P V P T P S P P F T K G V  
 CTC AGC AGT GGG GCG CCC CCA GTG CCC ACT CCT TCC CCA CCT CCT TTT ACT AAG GGT GTG  
 4924 4933 4942 4951 4960 4969  
  
 V T D S K V T S A F Q M T S N R V V T I  
 GTT ACA GAC AGC AAA GTC ACA TCA GCT TTC CAG ATG AGG TCA ANT AGA GTG GTC ACC ATA  
 4984 4993 5002 5011 5020 5029  
  
 Y E S S R H N T D L Q Q P S A E A S P N  
 TAT GAA TCT TCA AGG CAC ANT ACA GAT CTG CAG CAA CCC TCA GCA GAG GCT AGC CCC AAT  
 5044 5053 5062 5071 5080 5089  
  
 P E I I T Q T T D S P S N L F P S T S V  
 CCT GAG ATC ATA ACT GGA ACC ACT GAC TCT CCC TCT AAT CTG TTT CCA TCC ACT TCT GTG  
 5104 5113 5122 5131 5140 5149  
  
 P A L R V D K P Q N S K W K P S P W P E  
 CCA GCA CTA AGG GTA GAT AAA CCA CAG AAT TCT AAA TGG AAG CCC TCT CCC TGG CCA GAA  
 5164 5173 5182 5191 5200 5209

FIG. 3Q

H K Y Q L K S Y S E T I E K G K R P A V  
 CAC AAA TAT CAG CTC AAG TCA TAC TCC GAA ACC ATT GAG AAG GGC AAA AGG CCA GCA GTA  
 5224 5233 5242 5251 5260 5269

S M S P H L S L P E A S T H A S H W N T  
 AGC ATG TCC CCC CAC CTC AGC CTT CCA GAG GCC AGC ACT CAT GCC TCA CAC TGG AAT ACA  
 5284 5293 5302 5311 5320 5329

Q K H A E K S V F D K K P G Q N P T S K  
 CAG AAG CAT GCA GAA AAG AGT GTT TTT GAT AAG AAA CCT GGT CAA AAC CCA ACT TCC AAA  
 5344 5353 5362 5371 5380 5389

H L P Y V S L P K T L L K K P R I I G G  
 CAT CTG CCT TRC GTC TCT CTA CCT AAG ACT CTA TTG AAA AAG CCA AGA ATA ATT GGA GGA  
 5404 5413 5422 5431 5440 5449

K A A S F T V P A N S D V F L P C E A V  
 AAG GCT GCA AGC TTT ACA GTT CCA GCT AAT TCA GAC GTT TTT CTT TGT GAG GCT GTT  
 5464 5473 5482 5491 5500 5509

G D P L P I I H W T R V S S G X E I S Q  
 GGA GAC CCA CTG CCC ATC ATC CAC TGG ACC AGA GTT TCA TCA GGA NTT GAA ATA TCC CAA  
 5524 5533 5542 5551 5560 5569



## FIG. 3R

G T Q K S R F H V L P N G T L S I Q R V  
 GGG ACA CAG AAA AGC CGG TTC CAC GTG CTT CCC AAT GGC ACC TTG TCC ATC CAG AGG GTC  
 5584 5593 5602 5611 5620 5629

S I Q D R G Q Y L C S A F N P L G V D H  
 AGT ATT CAG GAC GGC GGA CAG TAC CTG TGC TCT GCA TTT AAT CCA CTG GGC GTA GAC CAT  
 5644 5653 5662 5671 5680 5689

F H V S L S V V F Y P A R I L D R H V K  
 TTT CAT GTC TCT TTG TCT GTG GTT TTT TAC CCG GCA AGG ATT TTG GAC AGA CAT GTC AAG  
 5704 5713 5722 5731 5740 5749

E I T V H F G S T V E L K C R V E G M P  
 GAG ATC ACA GTT CAC TTT GGA AGT ACT GTG GAA CTA RAG TGC AGA GTG GAG GGT ATG CCG  
 5764 5773 5782 5791 5800 5809

R P T V S W I L A N Q T V V S E T A K G  
 AGG CCT ACG GTT TCC TGG ATA CTT GCA AAC CAA ACG GTG GTC TCA GAA ACG GCC AAG GGA  
 5824 5833 5842 5851 5860 5869

S R K V W V T P D G T L I I Y N L S L Y  
 AGC AGA AAG GTC TGG GGA ACA CCT GAT GGA ACA TTG ATC ATC TAT AAT CTG AGT CTT TAT  
 5884 5893 5902 5911 5920 5929

## FIG. 3S

D R G F Y K C V A S N P S G Q D S L L V  
 GAT CGT GGT TTT TAC RAG TGT GTG GCC AGC RAC CCA TCT GGC CAG GAT TCA CTG TTG GTT 5944  
 5944 5953 5962 5971 5980 5989

K I Q V I T A P P V I I E Q K R Q A I V  
 AAG ATA CAA GTC ATC ACA GCT CCC CCT GTC ATT ATA GAG CAA AAG AGG CAA GCC ATC GTT 6004  
 6004 6013 6022 6031 6040 6049

G V L G G S L K L P C T A K G T P Q P S  
 GGG GTT TTA GGT GGA AGT TTG AAA CTG CCC TGC ACT GCA AAA GGA ACT CCC CAG CCT AGT 6064  
 6064 6073 6082 6091 6100 6109

V H W V L Y D G T E L K P L Q L T H S R  
 GTT CAC TGG GTC CTT TAT GAT GGG ACT GAA CTA AAA CCA TTG CAG TTG ACT CAT TCC AGA 6124  
 6124 6133 6142 6151 6160 6169

F F L Y P N G T L Y I R S I A P S V R G  
 TTT TTC TTG TAT CCA RAG GAT GGA ACT CTG TAT ATA AGA AGC ATC GCT CCT TCA GTG AGG GGC 6184  
 6184 6193 6202 6211 6220 6229

T Y E C I A T S S S G S E R R V V I L T  
 ACT TAT GAG TGC ATT GCC ACC AGC TCC TCA GGC TCA GAG AGA AGG GTA GTG ATT CTT ACT 6244  
 6244 6253 6262 6271 6280 6289

## FIG. 3T

V E E G E T I P R I E T A S Q K W T E V  
 GTG GAA GAG GGA ACA ATC CCC AGG ATA GAA ACT GCC TCT CAG AAA TGG ACT GAG GTG  
 6304 6313 6322 6340 6349

N L G E K L L L N C S A T G D P K P R I  
 AAT TTG GGT GAG AAA TTA CTA CTG AAC TGC TCA GGT ACT GGG GAT CCA AAG CCT AGA ATA  
 6364 6373 6382 6391 6400 6409

I W R L P S K A V I D Q W H R M G S R I  
 ATC TGG AGG CTG CCA TTC CCA TCC TAG GCT GTC ATC GAC CAG TGG CAC AGA ATG GGC AGC CGA ATC  
 6424 6433 6442 6451 6460 6469

H V Y P N G S L V V G S V T E K D A G D  
 CAC GTC TAC CCA AAT GGA TCC TTG GTG GTT GGG TCA GTG ACG GAA AAA GAC GCT GGT GAC  
 6484 6493 6502 6511 6520 6529

Y L C V A R N K M G D D L V L M H V R L  
 TAC TTA TGT GTG GCA AGA AAC AAA ATG GGA GAT GAC CTA GTC CTG ATG CAT GTC CGC CTG  
 6544 6553 6562 6571 6580 6589

R L T P A K I E Q K Q Y F K K Q V L H G  
 AGA TTG ACA OCT GCC AAA ATT GAA CAG AAG CAG TAT TTT TAG AAG CAA GTG CTC CAT GGG  
 6604 6613 6622 6631 6640 6649

FIG. 3U

K D F Q V D C K A S G S P V P E V S W S  
 AAA GAT TTC CAA GTT GAC TGC AAG GCC TCT GGC TCC CCT GTG GAG GTA GTC TGG AGT  
 6664 6724 6733 6682 6691 6700 6709  
  
 L P D G T V L N N V A Q A D D S G Y R T  
 TTG CCT GAT GGG ACA GTG CTC AAC AAT GTA GCC CAA GCT GAT GGC AGT GGC TAT AGG ACC  
 6724 6733 6742 6751 6760 6769  
  
 K R Y T L F H N G T L Y F N N V G M A E  
 AAG AGG TAC ACC CTT TTC CAC AAT GGA ACC TTG TAT TTC AAC AAC AAT GGT GGG ATG GCA GAG  
 6784 6793 6802 6811 6820 6829  
  
 E G D Y I C S A Q N T L G K D E M K V H  
 GAA GGA GAT TAT ATC TGC TCT GCC CAG AAC ACC TTA GGG AAA GAT GAG ATG AAA GTC CAC  
 6844 6853 6862 6871 6880 6889  
  
 L T V L T A I P R I R Q S Y K T T M R L  
 CTA ACA GTT CTA ACA GGC ATC CCA CGG ATA AGG CAA AGC TAC AAG ACC ACC ATG AGG CTC  
 6904 6913 6922 6931 6940 6949  
  
 R A G E T A V L D C E V T G E P K P N V  
 AGG GCT GGA GRA ACA GGT GTC CTT GAC TGC GAG GTC ACT GGG GRA CCG AAG CCC AAT GTA  
 6964 6973 6982 6991 7000 7009

FIG. 3V

F	W	L	P	S	N	V	I	S	F	S	N	D	R	F	T	F	H
TTT	TGG	TIG	CTG	CCT	TCC	AAC	AAT	GTC	ATT	TCA	TTT	TCC	AAT	GAC	AGG	TTC	ACA
7024				7033				7042					7051			7060	7069
A	N	R	T	L	S	I	H	K	V	K	P	L	D	S	G	D	Y
GCC	AAT	AGA	ACT	TIG	TCC	ATC	CAT	AAA	GTG	AAA	CCA	CTT	GAC	TCT	GGG	GAC	TAT
7084				7093				7102				7111			7120		7129
V	A	Q	N	P	S	G	D	T	K	T	Y	K	L	D	I	V	S
GTA	GCT	CAG	AAT	CCT	AGT	GGG	GAT	GAC	ACT	AAG	ACA	TAC	AAA	CTG	GAC	ATT	GTC
7144				7153				7162				7171			7180		7189
P	P	L	I	N	G	L	Y	A	N	K	T	V	I	K	A	T	A
CCT	CCA	TTA	ATC	AAT	GGC	CTG	TAT	GCA	AAC	AAG	ACT	GTT	ATT	AAA	GCC	ACA	GCC
7204				7213				7222				7231			7240		7249
H	S	K	K	Y	F	D	C	R	A	D	G	I	P	S	S	Q	V
CAC	TCC	AAA	AAA	TAC	TTT	GAC	TGC	AGA	GCA	GAT	GGG	ATC	CCA	TCT	TCC	CAG	GTC
7264				7273				7282				7291			7300		7309
I	M	P	G	N	I	F	L	P	A	P	Y	F	G	S	R	V	T
ATT	ATG	CCA	GGC	AAT	ATT	TTT	CTC	CCA	GCT	CCA	TAC	TTT	GGA	AGC	AGA	GTC	ACG
7324				7333				7342				7351			7360		7369

## FIG. 3W

P N G T L E M R N I R L S D S A D F T C  
 CCA AAT GGA ACC TTG GAG ATG ACG AAC ATC CGG CTT TCT GAC TCT GCG GAC TTC ACC TGT  
 7384 7393 7402 7411 7420 7429  
  
 V V R S E G G E S V L V V Q L E V L E M  
 GTG GTT CGG AGC GAG GGA GGA GAG AGT GTG TTG GTA GTG CAG TTA GAA GTC CTA GAA ATG  
 7444 7453 7462 7471 7480 7489  
  
 L R R P T F R N P F N E K V I A Q A G K  
 CTG AGA AGA CCA ACA TTC AGA AAC CCA TTC AAC GAA AAA GTC ATC GCC CAA GCT GGC AAG  
 7504 7513 7522 7531 7540 7549  
  
 P V A L N C S V D G N P P P E I T W I L  
 CCC GTA GCA CTG AAC TGC TCT GTG GAT GGG AAC CCC CCA CCT GAA ATT ACC TGG ATC TTA  
 7564 7573 7582 7591 7600 7609  
  
 P D G T Q F A N R P H N S P Y L M A G N  
 CCT GAC GGC ACA CAG TTT GCT AAC AGA CCA CAC AAT TCC CCG TAT CTG ATG GCA GGC AAT  
 7624 7633 7642 7651 7660 7669  
  
 G S L I L Y K A T R N K S G K Y R C A A  
 GGC TCT CTC ATC CTT TAC AAA GCA ACT CGG AAC AAG TCA GGG AAG TAT CGC TGT GCA GCC  
 7684 7693 7702 7711 7720 7729

FIG. 3X

R N K V G Y I E K L I L L E I G Q K P V  
 AGG AAT AAG GTT GGC TAC ATC GAG AAA CTC ATC CTG TTA GAG ATT GGG CAG AAG CCA GTC  
 7744 7753 7762 7771 7780 7789  
  
 I L T Y E P G M V K S V S G E P L S L H  
 ATT CTG ACA TAC CAA CCA GCG ATG GTG AAG AGC GTC AGT GGG GAA CCG TTA TCA CTG CAT  
 7804 7813 7822 7831 7840 7849  
  
 C V S D G I P K P N V K W T T P G G H V  
 TGT GTG TCT GAT GGG ATC CCC AAG CCA AAT GTC AAG TGG ACT ACA CCG GGT GGC CAT GTA  
 7864 7873 7882 7891 7900 7909  
  
 I D R P Q V D G K Y I L H E N G T L V I  
 ATC GAC AGG CCT CAA GTG GAT GGA AAA TAC ATA CTG CAT GAA AAT GGC ACG CTG GTC ATC  
 7924 7933 7942 7951 7960 7969  
  
 K A T T A H D Q G N Y I C R A Q N S V G  
 AAA GCA ACA ACA GCT CAC GAC CAA GGA AAT TAT ATC TGT AGG CCT CAA AAC AGT GTT GGC  
 7984 7993 8002 8011 8020 8029  
  
 Q A V I S V S V M V A Y P P R I I N Y  
 CAG GCA GTT ATT AGC GTG TCA GTG ATG GTT GTG GGC TAC CCT CCC CGA ATC ATA AAC TAC  
 8044 8053 8062 8071 8080 8089

FIG. 3Y

L P R N M L R R T G E A M Q L H C V A L  
 CTA CCC AGG AAC ATG CTC AGG ACA GGG GAA GCC ATG CAC CTC CAC TGT GTG GCC TTG  
 8104 8113 8122 8131 8140 8149

G I P K P K V T W E T P R H S L L S K A  
 GGA ATC CCC AAG CCA AAA GTC ACC TGG GAG ACG CAC AGA CAC TCC CTG CTC TCA AAA GCA  
 8164 8173 8182 8191 8200 8209

T A R K P H R S E M L H F Q G T L V I Q  
 ACA GCA AGA AAA CCC CAT AGA AGT GAG ATG CTT CAC CCA CAA GGT ACG CTG GTC ATT CAG  
 8224 8232 8242 8251 8260 8269

N L Q T S D S G V Y K C R A Q N L L G T  
 AAT CTC CAA ACC TCG GAT TCC GGA GTC TAT AAG TGC AGA GCT CAG AAC CTA CTT GGG ACT  
 8284 8293 8302 8311 8320 8329

D Y A T T Y I Q V L  
 GAT TAC GCA ACA ACT TAC ATC CAG GTA CTC TGA CAG GAA GGG GGA GAC TAA AAT TCA ACA  
 8344 8353 8362 8371 8380 8389



## FIG. 3Z

GAA GTC CAC AIC CAC AGG GTT 'TAT TTT TTG GAA GAA GTT TAA TCA AAG GCA GCC ATA GGC	8413	8422	8431	8440	8449
8404					
ATG TAA ATG AGT CTG AAT ACA TTT ACA GTA TTA AAT TTA CAA TGG ACA TGC GAT GAG ACT	8473	8482	8491	8500	8509
8464					
TGT AAA TGA AAG CAT TGT GAA CTG AAA CCG AGT CTG TGT GGA TCT CAA AGC AAA CTC TTA	8533	8542	8551	8560	8569
8524					
ACT TAA GGC ACT TTG ATT TTG CCA ACA AAT AAT AAC AAA CAT TAA GAG AAA AAA ATG ATC	8593	8602	8611	8620	8629
8584					
CAC TAC GAA ATA ACA AAC GGC TAA TGC ACC TGA ATT CTC AGT AAA AAG ACC TTT CTC TCG	8653	8662	8671	8680	8689
8644					

## FIG. 3AA

CTA ACA GGT GCC AGC TGC CTC GTG TCT GTT TCC TAC CAA TGT CAC AAA CAT CGC ACA CAG  
8704 8713 8722 8731 8740 8749

GGT GAA TGG AGT CAA CGG GAA AGA TTA AGT TTG CCG TCT GTG TAA ATC TCA ATG TAC AAA  
8764 8773 8782 8791 8800 8809

TAT TCT GTC NCT GGT TTA TAA ACA TTT TGA TAA AAC CGA AAA AAA AAA AAA AAA AAA  
8824 8833 8842 8851 8860 8869

AAA AAA AA

608-Rat

PROBE: 5' fragment of 608(Rat)Target: (mRNA)

N-normal

T-Mechanical force

N T

28S

18S →

*FIG. 4*

## FIG. 5A

Length of 608  
Translated from: 71 to: 762 (Entire region);  
Genetic Code used: Universal;

7680 bp; Listed from: 2 to: 7680;

```

Frame 2  L   E   S   L   Y   L   M   G   N   P   W   T   C   D   C   H   L   K
          CTA GAA AGC CTG TAT TTG CAT GGA AAC CCA TGG ACC TGT GAC TGC CAT TTA AAG
          10      19      28      37      46      55

W   L   S   E   W   M   Q   G   N   P   D   I   I   K   C   K   K   D   R   S
TGG TTG TCT GAG TGG ATG CAG GGA AAC CCA GAT ATA ATA AAA TGC AAG AAA GAC AGA AGC
          64      73      82      91      100      109

S   S   S   P   Q   Q   C   P   L   C   M   N   P   R   I   S   K   G   R   P
TCT TCC AGT CCT CAG CAA TGT CCC CTT TGC ATG AAC CCC AGG ATC TCT AAA GGC AGA CCC
          124      133      142      151      160      169

F   A   M   V   P   S   G   A   F   L   C   T   K   P   T   I   D   P   S   L
TTT GCT ATG GTA CCA TCT GGA GCT TTC CTA TGT ACA AAG CCA ACC ATT GAT CCA TCA CTG
          184      193      202      211      220      229

```

## FIG. 5B

K	S	K	S	L	V	T	Q	E	D	N	G	S	A	S	T	S	P	Q	D
AAG	TCA	AAG	AGC	GTG	GTT	ACT	CAG	GAG	GAC	AAT	GGA	TCT	GAC	TCC	ACC	TCA	CCT	CAA	GAT
244		244		253		253		262		262		271		280		280		289	
F	I	E	P	F	G	S	L	S	L	N	M	T	D	L	S	G	N	K	A
TTC	ATA	GAA	CCC	TTT	GGC	TCC	TTG	TCT	TTG	AAC	ATG	ACA	GAC	CTG	TCT	GGA	AAT	AAG	GCC
304		304		313		313		322		331		331		340		349		349	
D	M	V	C	S	I	Q	K	P	S	R	T	S	P	T	A	F	T	E	E
GAC	ATG	GTC	TGT	AGT	ATC	CAA	AAG	CCA	TCA	AGG	ACA	TCA	CCA	ACT	GCA	TTC	ACT	GAA	GAA
364		364		373		373		382		391		391		400		409		409	
N	D	Y	I	M	L	N	A	S	F	S	T	N	L	V	C	S	V	D	Y
AAT	GAC	TAC	ATC	ATG	CTA	AAT	GCG	TCA	TTT	TCC	ACA	AAT	CTT	GTG	TGC	AGT	GTA	GAT	TAT
424		424		433		433		442		451		451		460		469		469	
N	H	I	Q	P	V	W	Q	L	L	A	L	Y	S	D	S	P	L	I	L
AAT	CAC	ATC	CAG	CCA	GTG	TGG	CAA	CTT	CTG	GCT	TTA	TAC	AGT	GAC	TCT	CCT	CTG	ATA	CTA
484		484		493		493		502		511		511		520		529		529	

## FIG. 5C

E	R	K	P	Q	L	T	E	T	P	S	L	S	S	R	Y	K	Q	V	A
GAA	AGG	AAG	CCC	CAG	CTT	ACC	GAG	ACT	CCT	TCA	CTG	TCT	TCT	AGA	TAT	AAA	CAG	GTG	GCT
			544		553				562		571			580					589
L	R	P	E	D	I	F	T	S	I	E	A	D	V	R	A	D	P	F	W
CTT	AGG	CCT	GAA	GAC	ATT	TTT	ACC	AGC	ATA	GAG	GCT	GAT	GTC	AGA	GCA	GAC	CCT	TTT	TGG
			604		613				622		631			640					649
F	Q	Q	E	K	I	V	L	Q	L	N	R	T	A	T	T	L	S	T	L
TTC	CAA	CAA	GAA	AAA	ATT	GTC	TTG	CAG	CTG	AAC	AGA	ACT	GCC	ACC	ACA	CTT	AGC	ACA	TTA
			664		673				682		691			700					709
Q	I	Q	F	S	T	D	A	Q	I	A	L	P	R	A	E	M	R	A	E
CAG	ATC	CAG	TTT	TCC	ACT	GAT	GCT	CAA	ATC	GCT	TTA	CCA	AGG	GCG	GAG	ATG	AGA	GCG	GAG
			724		733				742		751			760					769
R	L	K	W	T	M	I	L	M	M	N	N	P	K	L	E	R	T	V	L
AGA	CTC	AAA	TGG	ACC	ATG	ATC	CTG	ATG	ATG	AAC	AAT	CCC	AAA	CTG	GAA	CGC	ACT	GTC	CTG
			784		793				802		811			820					829

## FIG. 5D

V G G T I A L S C P G K G D P S P H L E  
 GTT GGC GGC ACT ATT GCC CTG AGC TGT CCA GGC AAA GGC GAC CCT TCA CCT CAC TTG GAA  
 844 853 862 871 880 889

W L L A D G S K V R A P Y V S E D G R I  
 TGG CTT CTA GCT GAT GGG AGT AAA GTG AGA GCC CCT TAC GTT AGC GAG GAT GGG CGA ATC  
 904 913 922 931 940 949

L I D K N G K L E L Q M A D S F D A G L  
 CTA ATA GAC AAA AAT GGG AAG TTG GAA CTG CAG ATG GCT GAC AGC TTT GAT GCA GGT CTT  
 964 973 982 991 1000 1009

Y H C I S T N D A D A D V L T Y R I T V  
 TAC CAC TGC ATA AGC ACC AAT GAT GCA GAT GCG GAT GTT CTC ACA TAC AGG ATA ACT GTG  
 1024 1033 1042 1051 1060 1069

V E P Y C E S T H D S G V Q H T V T G  
 GTA GAG CCC TAT GGA GAA AGC ACA CAT GAC AGT GGA GTC CAG CAC ACA GTG GTT ACG GGT  
 1084 1093 1102 1111 1120 1129

## FIG. 5E

E	T	L	D	L	P	C	L	S	T	G	V	P	D	A	S	I	S	W	I
GAG	ACG	CTC	GAC	CTT	CCA	TGC	CTT	TCC	ACG	GGT	GTT	CCA	GAT	GCT	TCT	ATT	AGC	TGG	ATT
		1144			1153			1162			1171			1180					1189
L	P	G	N	T	V	F	S	Q	P	S	R	D	R	Q	I	L	N	N	G
CTT	CCA	GGG	AAC	ACT	GTG	TTC	TCT	CAG	CCA	TCA	AGA	GAC	AGG	CAA	ATT	CTT	RAC	AAT	GGG
		1204			1213			1222			1231			1240					1249
T	L	R	I	L	Q	V	T	P	K	D	Q	G	H	Y	Q	C	V	A	A
ACC	TTA	AGA	ATA	TTA	CAG	GTT	ACG	CCA	AAA	GAT	CAA	GGT	CAT	TAC	CAA	TGT	GTG	GCT	GCC
		1264			1273			1282			1291			1300					1309
N	P	S	G	A	D	F	S	S	F	K	V	S	V	Q	K	K	G	Q	R
RAC	CCA	TCA	GGG	GCC	GAC	TTT	TCC	AGT	TTT	AAA	GTT	TCA	GTT	CAA	AAG	AAA	GGC	CAA	AGG
		1324			1333			1342			1351			1360					1369
M	V	E	H	D	R	E	A	G	G	S	G	L	G	E	P	N	S	S	V
ATG	GTT	GAG	CAT	GAC	AGG	GAG	GCA	GGT	GGA	TCT	GGA	CTT	GGA	GAA	CCC	AAC	TCC	AGT	GTT
		1384			1393			1402			1411			1420					1429



## FIG. 5F

S L K Q P A S L K L S A S A L T G S E A  
 TCC CTT AAG CAG CCA GCA TCT TIG AAA CTC TCT GCA TCA GCT TIG ACA GGG TCA GAG GCT  
 1444 1453 1462 1471 1480 1489  
  
 G K Q V S G V H R K N K H R D L I H R R  
 GGA AAA CAA GTC TCC GGT GTA CAT AGG AAG AAC AAA CAT AGA GAC TTA ATA CAT CGG CGG  
 1504 1513 1522 1531 1540 1549  
  
 R G D S T L R R F R E H R R Q L P L S A  
 CGT GGG GAT TCC ACG CTC CGG CGA TTC AGG GAG CAT AGG AGG CAG CTC CCT CTC TCT GCT  
 1564 1573 1582 1591 1600 1609  
  
 R R I D P Q R W A A L L E K A K K N S V  
 CGG AGA ATT GAC CCG CAA CGC TGG GCA GCA CTT CTA GAA AAA GCC AAA AAG AAT TCT GTG  
 1624 1633 1642 1651 1660 1669  
  
 P K K Q E N T T V K P V P L A V P L V E  
 CCA AAA AAG CAA GAA AAT ACC ACA GTA AAG CCA GTG CCA CTG GCT GTT CCC CTC GTG GAA  
 1684 1693 1702 1711 1720 1729

## FIG. 5G

L T D E E K D A S G M I P P D E E F M V	
CTC ACT GAC GAG GAA AAG GAT GCC TCT GGC ATG ATT CCT CCA GAT GAA GAA TTC ATG GTT	1744 1753 1762 1771 1780 1789
L K T K A S G V P G R S P T A D S G P V	
CTG AAA ACT AAG GCT TCT GGT GTC CCA GGA AGG TCA CCA ACT GCT GAC TCT GGA CCA GTA	1804 1813 1822 1831 1840 1849
N H G F M T S I A S G T E V S T V N P Q	
AAT CAT GGT TTT ATG ACG AGT ATA GCT TCT GGC ACA GAA GTC TCA ACT GTG AAT CCA CAA	1864 1873 1882 1891 1900 1909
T L Q S E H L P D F K L F S V T N G T A	
ACA CTA CAA TCT GAG CAC CTT CCT GAT TTC AAA TTA TTT AGT GTA ACA AAC GGT ACA GCT	1924 1933 1942 1951 1960 1969
V T K S M N P S I A S K I E D T T N Q N	
GTG ACA AAG AGT ATG AAC CCA TCC ATA GCA AGC AAA ATA GAA GAT ACA ACC AAC CAA AAC	1984 1993 2002 2011 2020 2029

## FIG. 5H

P I I I F P S V A E I R D S A Q A G R A	2089
CCA ATC ATT ATC TTT CCA TCA GTA GCT GAA ATT CGA GAT TCT GCT CAG GCA GGA AGA GCA	2080
	2071
	2062
	2053
	2044
S S Q S A H P V T G G GGA AAC ATG GCT ACC TAT GGC CAT ACC AAC	2149
TCT TCC CAA AGT GCA CAC CCT GTA ACA GGG GGA AAC ATG GCT ACC TAT GGC CAT ACC AAC	2140
	2131
	2122
	2113
	2104
T Y S S F T S K A S T V L Q P I N P T E	2209
ACA TAT AGT AGC TTT ACC AGC AAA GCC AGT ACA GTC TTG CAG CCA ATA AAT CCA ACA GAA	2200
	2191
	2182
	2173
S Y G P Q I P I T G V S R P S S D I S	2269
AGT TAT GGA CCT CAG ATA CCT ATT ACA GGA GTC AGC AGA CCT AGC AGT AGT GAC ATC TCT	2260
	2251
	2242
	2233
	2224
S H T T A D P S F S S H P S G S H T T A	2329
TCT CAC ACT ACT GCA GAC CCT AGC TTC TCC AGT GAC CCT TCA GGT TCA CAC ACC ACT GGC	2320
	2311
	2302
	2293
	2284

## FIG. 5I

S S L F H I P R N N N T G N F P L S R H  
 TCG TCT TTA TTT CAC ATT CCT AGA AAC AAC AAT ACA GGT AAC TTC CCC TTG TCC AGG CAC  
 2344 2353 2362 2371 2380 2389  
  
 L G R E R T I W S R G R V K N P H R T P  
 TTG GGA AGA GAG AGG ACA ATT TCG AGC AGA GGG AGA GTT AAA AAC CCA CAT AGA ACC CCA  
 2404 2413 2422 2431 2440 2449  
  
 V L R R H R H R T V R P A I K G P A N K  
 GTT CTC CGA CGG CAT AGA CAC AGG ACT GTG AGG CCA GCA ATC AAG GGA CCT GCT AAC AAA  
 2464 2473 2482 2491 2500 2509  
  
 N V S Q V P A T E Y P G M C H T C P S A  
 AAT GTG AGC CAA GTT CCA GCC ACA GAG TAC CCT GGG ATG TGC CAC ACA TGT CCT TCC GCA  
 2524 2533 2542 2551 2560 2569  
  
 E G L T V A T A A L S V P S S H S A L  
 GAG GGG CTC ACA GTG GCT ACT GCA GCA CTG TCA GTT CCA AGT TCA TCC CAC AGT GCC CTC  
 2584 2593 2602 2611 2620 2629

## FIG. 5J

P K T N N V G V I A E E S T T V V K K P	
CCC AAA ACT AAT GTT GGG GTC ATA GCA GAA GAG TCT ACC ACT GTG GTC AAG AAA CCA	2689
	2644
	2653
	2662
	2671
	2680
	2690
	2704
L L F K D K Q N V D I E I I T T T K	
CTG TTA CTA TTT AAG GAC AAA CAA AAT GTA GAT AIT GAG ATA ATA ACA ACC ACT ACA AAA	2749
	2704
	2713
	2722
	2731
	2740
	2750
	2764
	2773
	2782
	2791
	2800
	2809
	2818
	2824
	2833
	2842
	2851
	2860
	2869
	2878
	2884
	2893
	2902
	2911
	2920
	2929

## FIG. 5K

P S I P T S T K F S K R K T P L H Q I F  
 CCC TCA ATA CCA ACA AGC ACA AAA TTC TCA AAG AGG AAA ACT CCC TTG CAC CAG ATC TTT  
 2944 2953 2962 2971 2980 2989  
  
 V N N Q K K E G M L K N P Y Q F G L Q K  
 GTA AAT AAC CAG AAG CAG GGG ATG TTA AAG AAT CCA TAT CAA TTC GGT TTA CAA AAG  
 3004 3013 3022 3031 3040 3049  
  
 N P A A K L P K I A P L L P T G Q S S P  
 AAC CCA GCC GCA AAG CTT CCC AAA ATA GCT CCT CTT TTA CCC ACA GGT CAG AGT TCC CCC  
 3064 3073 3082 3091 3100 3109  
  
 S D S T T L L T S P P P A L S T T M A A  
 TCA GAT TCT ACA ACT CTC TTG ACA AGT CCG CCA CCA GCT CTG TCT ACA ACA ATG GCT GCC  
 3124 3133 3142 3151 3160 3169  
  
 T Q N K G T E V V S G A R S L S A G K K  
 ACT CAG AAC AAG GGC ACT GAA GTA GTA TCA GGT GCC AGA AGT CTC TCA GCA GGG AAG AAG  
 3184 3193 3202 3211 3220 3229

## FIG. 5L

Q P F T N S S P V L P S T I S K R S N T  
 CAG CCC TTC ACC AAC TCC TCT TCT CCA GTG CTT CCT AGC ACC ATA AGC AAG AGA TCT AAT ACA  
 3244 3253 3262 3271 3280 3289  
  
 L N F L S T E T P T V T S P T A T A S V  
 TTA AAC TTC TTG TCA ACG GAA ACC CCC ACA GTG ACA AGT CCT ACT GCT ACT GCA TCT GTC  
 3304 3313 3322 3331 3340 3349  
  
 I M S E T Q R T R S K E A K D Q I K G P  
 ATT ATG TCT GAA ACC CAA CGA ACA AGA TCC AAA GAA GCA AAA GAC CAA ATA AAG GGG CCT  
 3364 3373 3382 3391 3400 3409  
  
 R K N R N N A N T T P R Q V S G Y S A Y  
 CGG AAG AAC AGA AAC AAC GCA AAC ACC ACC CCC AGG CAG GTT TCT GGC TAT AGT GCA TAC  
 3424 3433 3442 3451 3460 3469  
  
 S A L T A D T P L A F S H S P R Q D D  
 TCA GCT CTA ACA ACA GCT GAT ACC CCC TTG GCT TTC AGT CAT TCC CCA CGA CAA GAT GAT  
 3484 3493 3502 3511 3520 3529

## FIG. 5M

G G N V S A V A Y H S T S L L A I T E  
 GGT GGA AAT GTA AGT GCA GTT GCT TAT CAC TCA ACA ACC TCT CTG GCC ATA ACT GAA  
 3544 3553 3562 3571 3580 3589  
  
 L F E K Y T Q T L G N T T A L E T T L L  
 CTG TTT GAG AAG TAC ACC CAG ACT TTG GGA AAT ACA GCA GCT TTG GAA ACA ACG TTG TTG  
 3604 3613 3622 3631 3640 3649  
  
 S K S Q E S T T V K R A S D T P P L L  
 AGC AAA TCA CAG GAG AGT ACC ACA GTG AAA AGA GCC TCA GAC ACA CCA CCA CTC CTC  
 3664 3673 3682 3691 3700 3709  
  
 S S G A P P V P T P S P P P F T K G V V  
 AGC AGT GGG GCG CCC CCA GTG CCC ACT CCT TCC CCA CCT CCT TTT ACT AAG GGT GTG GTT  
 3724 3733 3742 3751 3760 3769  
  
 T D S K V T S A F Q M T S N R V V T I Y  
 ACA GAC AGC AAA GTC ACA TCA GCT TTC CAG ATG ACG TCA AAT AGA GTG GTC ACC ATA TAT  
 3784 3793 3802 3811 3820 3829



## FIG. 5N

E S S R H N T D L Q Q P S A E A S P N P	
GAA TCT TCA AGG CAC AAT ACA GAT CTG CAG CAA CCC TCA TCA GAG GCT AGC CCC AAT CCT	3889
	3880
	3871
	3862
	3853
	3844
E I T G T T D S P S N L F P S T S V P	
GAG ATC ATA ACT GGA ACC ACT GAC TCT CCC TCT AAT CTG TTT CCA TCC ACT TCT GTG CCA	3949
	3940
	3931
	3922
	3913
	3904
A L R V D K P Q N S K W K P S P W P E H	
GCA CTA AGG GTA GAT AAA CCA CAG AAT TCT AAA TGG AAG CCC TCT CCC TGG CCA GAA CAC	4009
	4000
	3991
	3982
	3973
	3964
K Y Q L K S Y S E T I E K G K R P A V S	
AAA TAT CAG CTC AAG TCA TAC TCC GAA ACC ATT GAG AAG GGC AAA AGG CCA GCA GTA AGC	4069
	4060
	4051
	4042
	4033
	4024
M S P H L S L P E A S T H A S H W N T Q	
ATG TCC CCC CAC CTC AGC CTT CCA GAG GCC AGC ACT CAT GCC TCA CAC TGG AAT ACA CAG	4129
	4120
	4111
	4102
	4093
	4084

FIG. 50

K H A E K S V F D K K P G Q N P T S K H	
AAG CAT GCA GAA AAG AGT GTT TTT GAT AAG AAA CCT GGT CAA AAC CCA ACT TCC AAA CAT	4144
	4153
	4162
	4171
	4180
	4189
L P Y V S L P K T L L K K P R I I G G K	
CTG CCT TAC GTC TCT CTA CCT AAG ACT CTA TTG AAA AAG CCA AGA ATA ATT GGA GGA AAG	4204
	4213
	4222
	4231
	4240
	4249
A A S F T V P A N S D V F L P C E A V G	
GCT GCA AGC TTT ACA GTT CCA GCT AAT TCA GAC GTT TTT CCT CTT TGT GAG GCT GTT GGA	4264
	4273
	4282
	4291
	4300
	4309
D P L P I I H W T R V S S G X E I S Q G	
GAC CCA CTG CCC ATC ATC CAC TGG ACC AGA GTT TCA TCA GGA NTT GAA ATA TCC CAA GGG	4324
	4333
	4342
	4351
	4360
	4369
T Q K S R F H V L P N G T L S I Q R V S	
ACA CAG AAA AGC CGG TTC CAC GTG CTT CCC AAT GGC ACC TTG TCC ATC CAG AGG GTC AGT	4384
	4393
	4402
	4411
	4420
	4429

FIG. 5P

I	Q	D	R	G	Q	Y	L	C	S	A	F	N	P	L	G	V	D	H	F
ATT	CAG	GAC	CGI	GGA	CAG	TAC	CTG	TGC	TCT	GCA	TTT	AAT	CCA	CTC	GGC	GTA	GAC	CAT	TTT
	4144			4453			4462			4471			4480			4489			

H	V	S	L	S	V	V	F	Y	P	A	R	I	L	D	R	H	V	K	E
CAAT	GTC	TCT	TTG	TCT	GNG	GTT	TTT	TAC	CCG	GCA	AGG	ATT	TTG	GAC	AGA	CAT	GTC	AAG	GAG
4504			4513		4513		4522				4531		4540				4549		

I	T	V	H	F	G	S	T	V	E	L	K	C	R	V	E	G	M	P	R
ATC	ACA	GTT	CAC	TTT	GGA	AGT	ACT	GTG	GAA	CTA	AAG	TGC	AGA	GTG	GAG	GGT	ATG	CCG	AGG
		4564		4573		4582		4591		4600		4609							

P T V S W I L A N Q T V G S E T A K G S  
CCCT ACG GTT TCC TGG ATA CTT GCA AAC CAA ACG GTG GTC TTA GAA ACG GCC AAG GGA AGC  
4624 4633 4634 4642 4651 4660 4669

R	K	V	W	V	T	P	D	G	T	L	I	I	Y	N	L	S	L	Y	D
AGA	AAG	GTC	TGG	GTA	ACA	CCT	GAT	GGA	ACA	TTG	ATC	ATC	TAT	AT	CTG	AGT	CTT	TAT	GAT
	4684				4693		4702		4711		4711		4720		4720		4729		

## FIG.5Q

R G F Y K C V A S N P S G Q D S L L V K  
 CGT GGT TTT TAC AAG TGT GTG GCC AGC AAC CCA TCT GGC CAG GAT TCA CTG TTG GTT AAG  
 4744 4753 4762 4771 4780 4789  
  
 I Q V I T A P P V I I E Q K R Q A I V G  
 ATA CAA GTC ATC ACA GCT CCC CCT GTC ATT ATA GAG CAA AAG AGG CAA GCC ATC GTT GGG  
 4804 4813 4822 4831 4840 4849  
  
 V L G G S L K L P C T A K G T P Q P S V  
 GTT TTA GGT GGA AGT TTG AAA CTG CCC TGC ACT GCA AAA GGA ACT CCC CAG CCT AGT GTT  
 4864 4873 4882 4891 4900 4909  
  
 H W V L Y D G T E L K P L Q L T H S R F  
 CAC TGG GTC CTT TAT TAT GAT GGG ACT GAA CTA AAA CCA TTG CAG TTG ACT CAT TCC AGA TTT  
 4924 4933 4942 4951 4960 4969  
  
 F L Y P N G T L Y I R S I A P S V R G T  
 TTC TTG TAT CCA TAT GGA ACT CTG TAT ATA AGA AGC ATC GCT CCT TCA GTG AGG GGC ACT  
 4984 4993 5002 5011 5020 5029

## FIG. 5R

Y	E	C	I	A	T	S	S	G	S	E	R	R	V	V	I	L	T	V	
TAT	GAG	TGC	ATT	GCC	ACC	AGC	TCC	TCA	GGC	TCA	GAG	ACA	AGG	GTA	GTG	ATT	CTT	ACT	GTG
																			5089
																			5080
																			5071
																			5062
																			5053
																			5044
E	E	G	E	T	I	P	R	I	E	T	A	S	Q	K	W	T	E	V	N
GAA	GAG	GGA	GAG	ACA	ATC	CCC	AGG	ATA	GAA	ACT	GCC	TCT	CAG	AAA	TGG	ACT	GAG	GTG	AAT
																			5149
																			5140
																			5131
																			5122
																			5113
																			5104
L	G	E	K	L	L	L	N	C	S	A	T	G	D	P	K	P	R	I	I
TTG	GGT	GAG	AAA	TTA	CTA	CTG	AAC	TGC	TCA	GCT	ACT	GGG	GAT	CCA	AAG	CCT	AGA	ATA	ATC
																			5209
																			5200
																			5191
																			5182
																			5173
																			5164
W	R	L	P	S	K	A	V	I	D	Q	W	H	R	M	G	S	R	I	H
TGG	AGG	CTG	CCA	TCC	AAG	GCT	GTG	ATC	GAC	CAG	TGG	CAC	AGA	ATG	GGC	AGC	CGA	ATC	CAC
																			5269
																			5260
																			5242
																			5233
																			5224
V	Y	P	N	G	S	L	V	V	G	S	V	T	E	K	D	A	G	D	Y
GTC	TAC	CCA	AAT	GGA	TCC	TTG	GTG	GTT	GGG	TCA	GTG	ACG	GAA	AAA	GAC	GCT	GCT	GAC	TAC
																			5329
																			5320
																			5311
																			5302
																			5293
																			5284

## FIG. 5S

L C V A R N K M G D D L V L M H V R L R  
 TTA TGT GTG GCA AGA AAC AAA ATG GGA GAT GAC CTA GTC CTG ATG CAT GTC CGC CTG AGA  
 5344 5353 5362 5371 5380 5389

L T P A K I E Q K Q Y F K K Q V L H G K  
 TTG ACA CCT GCC AAA ATT GAA CAG AAG CAG TAT TTT AAG AAG CAA GTG CTC CAT GGG AAA  
 5404 5413 5422 5431 5440 5449

D F Q V D C K A S G S P V P E V S W S L  
 GAT TTC CAA GTT GAC TGC AAG GCC TCT GGC TCC CCT GTG CCT GAG GTA TCC TGG AGT TTG  
 5464 5473 5500 5491 5509

P D G T V L N N V A Q A D D S G Y R T K  
 CCT GAT GGG ACA GTG CTC AAC AAT GTA GCC CAA GCT GAT GAC AGT GGC TAT AGG ACC AAG  
 5524 5533 5542 5551 5560 5569

## FIG. 5T

R	Y	T	L	F	H	N	G	T	L	Y	F	N	N	V	G	M	A	E	E
AGG	TAC	ACC	CTT	TTC	CAC	AAT	GGA	ACC	TTC	TAT	TTC	AAC	AAC	GTT	GGG	ATG	GCA	GAG	GAA
			5584			5593		5602			5611			5620					5629
G	D	Y	I	C	S	A	Q	N	T	L	G	K	D	E	M	K	V	H	L
CGA	GAT	TAT	ATC	TGC	TCT	GCC	CAG	AAC	ACC	TTA	GGG	AAA	GAT	GAG	ATG	AAA	GTC	CAC	CTA
			5644			5653		5662			5671			5680					5689
T	V	L	T	A	I	P	R	I	R	Q	S	Y	K	T	T	M	R	L	R
ACA	GTT	CTA	ACA	GCC	ATC	CCA	CGG	ATA	AGG	CAA	AGC	TAC	RAG	ACC	ACC	ATG	AGG	CTC	AGG
			5704			5713		5722			5731			5740					5749
A	G	E	T	A	V	L	D	C	E	V	T	G	E	P	K	P	N	V	F
GCT	GGA	GAA	ACA	GCT	GTC	CTT	GAC	TGC	GAG	GTC	ACT	GGG	GAA	CCG	AAG	CCC	AAT	GTA	TTT
			5764			5773		5782			5791			5800					5809
W	L	L	P	S	N	N	V	I	S	F	S	N	D	R	F	T	F	H	A
TGG	TTG	CUG	CCT	TCC	AAC	AAT	GTC	ATT	TCA	TTC	TCC	AAT	GAC	AGG	TTC	ACA	TTT	CAT	GCC
			5824			5833		5842			5851			5860					5869
N	R	T	L	S	I	H	K	V	K	P	L	D	S	G	D	Y	V	C	V
AAT	AGA	ACT	TTG	TCC	ATC	CAT	AAA	GTG	AAA	CCA	CTT	GAC	TCT	GGG	GAC	TAT	GTG	TGC	GTA
			5884			5893		5902			5911			5920					5929

## FIG. 5U

A	Q	N	P	S	G	D	T	K	T	Y	K	L	D	I	V	S	K	P	
GCT	CAG	AAT	CCT	AGT	GGG	GAT	GAC	ACT	AAG	ACA	TAC	AAA	CTG	GAC	ATT	GTC	TCT	AAA	CCT
			5944		5953		5962		5971		5980								
P	L	I	N	G	L	Y	A	N	K	T	V	I	K	A	T	A	I	R	H
CCA	TTA	ATC	AAT	GGC	CTG	TAT	GCA	AAC	AAG	ACT	GTT	ATT	AAA	GCC	ACA	GCC	ATT	CGG	CAC
			6004		6013		6022		6031		6040								
S	K	K	Y	F	D	C	R	A	D	G	I	P	S	S	Q	V	T	W	I
TCC	AAA	AAA	TAC	TTT	GAC	TGC	AGA	GCA	GAT	GGG	ATC	CCA	TCT	TCC	CAG	GTC	ACG	TGG	ATT
			6064		6073		6082		6091		6100								
M	P	G	N	I	F	L	P	A	P	Y	F	G	S	R	V	T	V	H	P
ATG	CCA	GGC	AAT	ATT	TTT	CTC	CCA	GCT	CCA	TAC	TTT	GGA	AGC	AGA	GTC	ACG	GTC	CAT	CCA
			6124		6133		6142		6151		6160								
N	G	T	L	E	M	R	N	I	R	L	S	D	S	A	D	F	T	C	V
AAT	GGA	ACC	TTG	GAG	ATG	AGG	AAC	ATC	CGG	CTT	TCT	GAC	TCT	GGG	GAC	TTC	ACC	TGT	GTG
			6184		6193		6202		6211		6220								





## FIG. 5W

N K V G Y I E K L I L L E I G Q K P V I  
 AAT AAG GTT GGC TAC ATC GAG AAA CTC ATC CTG TTA GAG ATT GGG CAG AAG CCA GTC ATT  
 6544 6553 6562 6571 6580 6589  
  
 L T Y E P G M V K S V S G E P L S L H C  
 CTG ACA TAC GAA CCA GGG ATG GTG AAG AGC GTC AGT GGG GAA CCG TTA TCA CTG CAT TGT  
 6604 6613 6622 6631 6640 6649  
  
 V S D G I P K P N V K W T T P G G H V I  
 GTG TCT GAT GGG ATC CCC AAG CCA AAT GTC AAG TGG ACT-ACA CCG GGT GGC CAT GTA ATC  
 6664 6673 6682 6691 6700 6709  
  
 D R P Q V D G K Y I L H E N G T L V I K  
 GAC AGG CCT CAA GTG GAT GGA AAA TAC ATA CTG CAT GAA AAT GGC ACG CTG GTC ATC AAA  
 6724 6733 6742 6751 6760 6769  
  
 A T T A H D Q G N Y I C R A Q N S V G Q  
 GCA ACA ACA GCT CAC GAC CAA GGA AAT TAT ATC TGT AGG GCT CAA AAC AGT GTT GGC CAG  
 6784 6793 6802 6811 6820 6829

## FIG. 5X

A	V	I	S	V	S	V	M	V	V	A	Y	P	P	R	I	I	N	Y	L
GCA	GTT	ATT	AGC	GTG	TCA	GTG	ATG	GTT	GTG	GCC	TAC	CCT	CCC	CGA	ATC	ATA	AAC	TAC	CTA
	6844				6853				6862		6871			6880			6889		
P	R	N	M	L	R	R	T	G	E	A	M	Q	L	H	C	V	A	L	G
CCC	AGG	AAC	ATG	CTC	AGG	AGG	ACA	GGG	GAA	GCC	ATG	CAG	CTC	CAC	TGT	GTG	GCC	TTG	GGA
	6904			6913				6922			6931			6940			6949		
I	P	K	P	K	V	T	W	E	T	P	R	H	S	L	L	S	K	A	T
ATC	CCC	AAG	CCA	AAA	GTG	ACC	TGG	GAG	ACG	CCA	AGA	CAC	TCC	CTG	CTC	TCA	AAA	GCA	ACA
		6964			6973			6982			6991			7000			7009		
A	R	K	P	H	F	S	E	M	L	H	P	Q	G	T	L	V	I	Q	N
GCA	AGA	AAA	CCC	CAT	AGA	AGT	GAG	ATG	CTT	CAC	CAA	CCA	GGT	ACG	CTG	GTG	ATT	CAG	AAT
			7024		7033			7042			7051			7060			7069		
L	Q	T	S	D	S	G	V	Y	K	C	R	A	Q	N	L	L	G	T	D
CTC	CAA	ACC	TGG	GAT	TCC	GGA	GTC	TAT	AAG	TGC	AGA	GCT	CAG	AAC	CTA	CTT	GGG	ACT	GAT
		7084			7093			7102			7111			7120			7129		

71/183

## FIG. 5Y

[illegible]

72/183

## FIG. 5Z

TAC GAA ATA ACA AAC GGC TAA TGC ACC TGA ATT CTC AGT AAA AAG ACC TTT CTC TCG CTA  
 7444 7453 7462 7471 7480 7489

ACA GTT GCC AGC TGC CTC GTG TCT GTT TCC TAC CAA TGT CAC AAA CAT CGC ACA CAG GGT  
 7504 7513 7522 7531 7540 7549

GAA TGG AGT CAA CGG CAA AGA TTA AGT TTG CGG TCT GTG TAA ATC TCA ATG TAC AAA TAT  
 7564 7573 7582 7591 7600 7609

TCT GTC NCT GGT TTA TAA ACA TTT TGA TAA AAC CGA AAA AAA AAA AAA AAA AAA AAA  
 7624 7633 7642 7651 7660 7669

AAA AA

## FIG. 6A

## 608: CLUSTAL X (1.64b) Multiple Sequence Alignment

gi 1017427 emb CAA62189	PDQMPVYPPAJITPLQDVTSEGQPARFCRCVSGTDLKVSWYSKDKKIK
gi 3328186	-MGRSPSWLYGVIGIILLATT-----CSSVNDKND-----PTCKSS
608	-----
gi 1017427 emb CAA62189	PSRFFRMT--QFEDTYQLEIAEAYPEDEGTYTFVANNVGVQVSSANLSL
gi 3328186	LAFFVDTGSMEDDILQVREGRAK----IFKTVHAQREKLIYNIWVPF
608	-----
gi 1017427 emb CAA62189	EAPE\$--ILHERIEQEIEMEMKAPVIKRRKIEPLEVALGHLAKETCIQS
gi 3328186	HDPYIGEIINTTDSYFMHQLSKVYHGGGDCPEKTLTGILKALQISLPS
608	-----
gi 1017427 emb CAA62189	APNVRFQWFKAGREIYESDKCSIRSSKYISSLEILRTQVDCGEYTCCKAS
gi 3328186	SFIYVFTDARSKDYHLEDEVLNTIOEK-QSS--VVFVMTGDCGHRTHPGF
608	-----
gi 1017427 emb CAA62189	NEYGSVSCATLTVTEAYPPTEFLSRPKSLTTFVGKAAKFICTVGTGTPVIE
gi 3328186	RTYEKIAAAS-----FGQVFHLEKSDVSTVL---EYVNHAVKOKKVVH
608	-----

## FIG. 6B

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

TIWQKGAALSPSPNWRISDAENKHILELSNLTIQDRGVYSCAKNKFGA  
LMYERERGGTVSRNIPV----DKHLSLTISLSGDK-----DDSDNLDI  
-----

DIQDAELIIDKPHFIKELEFVQSAINKKVHLEQVDEDRKVTVTWSKDG  
VLRDEPGTVDKRLYSKEGGTIDIKNVKLIIRLKDPSEGVWTVNTN-SRLK  
-----

QKLPPGKDYKICFEDKIAITLEIPLAKDKDSGTVYVCTASNEAGSSSCSATV  
HTIRVFGHGAVDKYGFASRPLDRIELARPREVLNQDT-----YLLINM  
-----

TVREPPSFVKYKVDPSYIMLPGESARLHCKLKGSPVIOVTFKNNKELSES  
TGLIPPGTVGEID--LVVDYHG-----HSLYKAVASPHRT-NPNMYFAGPF  
-----

NTVVRMYFNSEAILDIIDNVKVEDSGSYSCAENVNDVCGSDSCSTEIVIKPPF  
VPPKGLFF---VRVQGYD---EDNYEFMRIAPTAIGS-----VIVGGPR  
-----

FIG. 6C

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

SFIKLEPADIVRGTNALLQCEVSGTGPFPEISWFKDKKOIRSSKKYRLFS  
AFWSPHQEFVGRDLN--LSCVESASAYTIYWKTGEDIIGGPLEYHNT  
-----

OKSLVLEIFSFNSADVGEYECWANEVKGCGCMATHLLKEPPTFVKKVD  
DTSVWTIPELSILK--DAGEYECRWISNNGNYSVKTRVETRESPPFIFGR  
-----

DLIALGGQTVTLQAAVRGSEPISVTWKGGQEVIREDGKIKMSFNGVAVL  
NVSVPLGEAAFLHCSTRSAGEVEIRWTRYGATVFNQPTERNPTNG--TL  
-----

IIPDVQISFSGKTYCIAENEAGSOTSVGELIVKEPAKIIERAEILQVTAG  
KIHVTRADAGVYECVARNAGCMSTRKMRDIMEPPSVKVTQDVYFNMR  
-----

DPATLEYTVAGTPELKPKWYKDGRLVASKKYRISFKNNVAQLKIFYSAEL  
EGVNLSCEAMGDPKEVHWYFKGRHLLNDYKYQVGQDS--KELYIRDATH  
-----



FIG. 6D

gi|11017427|emb|CAA62189|  
gi|3328186  
608

gi|11017427|emb|CAA62189|  
gi|3328186  
608

gi|11017427|emb|CAA62189|  
gi|3328186  
608

gi|11017427|emb|CAA62189|  
gi|3328186  
608

gi|11017427|emb|CAA62189|  
gi|3328186  
608

HDGQYFEISNEVGSSCETTFTVLDRDIAPFFTKPLRNVDSVVGTCR  
HDEGTVECRAMSQAG---QARTTDDIMLATPKVEIIQNKKMWCRGDRV  
-----

-LDCKIAGSLPMR-VSWFKDGKEIAASDRYIAFVEGTASLEIIRVDMMD  
SFECKTIRGKPHPKIRWFKNKGLIRPDDY-IKINEG--QLHMGAKDED  
-----

AGNFTCRATNSVGSKDSSGALIVQEPSPFVTKPGSKDVLPGSAVCLKSTF  
AGAYSCVGENMAGKDVQVANLSVGRVTIIESPHTVRVNIEROVTLQCLA  
-----

FVKEPAAFIKRLSDHSVEFGKSIILESTYTGTLPISTVTKKDGFNITSE  
VIIPFVIDGERREAVAVIEGFSESELFCDNSNLTG-VDVWOKDGLTINQDT  
LPLSARRIDPQRWRALLEKAKKNSVPKQENTTVKEVPLAVPLVELITDEE  
.....\*

KCN---IVTTEKTCILEILNSTKRDAQYSCIEIENEAGRDVCGALVSTIE  
LRGDSFIQIPSSGKKMSFELSKARKSDSGRYTCIVRNPAEGARKLDFEAVND  
KDAS--GMIPPFDEEFVLTAKSGVPGRSPTADSGFVNHGEMTSTASGTE  
.....\*

## FIG. 6E

gi|1017427|emb|CAA62189|  
gi|3328186  
608gi|1017427|emb|CAA62129|  
gi|3328186  
608

```
gi|1017427|emb|CAA62189|
gi|3328186
608
```

gi|1017427|emb|CAA62189|  
gi|3328186  
608gi|1017427|emb|CAA62189|  
gi|3328186  
608

PPPVVTELEP--LEAAVG-DSVSLQCQVAGTPELTVSWYKGBTKLRPTPE  
PPPSIDELSSANIQTIVPYYPEINCVWSGSPHPKVYWLFDKPLDPD-S  
VSVTVNPQTLOSEHLPDFKLEFSVTNGTAVTKSMNPISAKIEDTTNQNPII

YRVTFTNNVATLVNKVNINDSGEYTCCKAENSIGTASCKTVFRIQERQLP  
AAAYELTNNGETLKIVRSQVEHAGTYTCEAQNNGVKARKDFLVRVT--AP  
IFPSVAEIRDSAQAGRASSQSAHPVTGGNNMATYGHNTYTSFTSK----

PSFARQLKDIEQTVGLPVLTCRLNGSAPIQVCWYRDGVLRLDHNELQTS  
PHIFEKEREEVAVRGDMLLTCNAESSVPLSSYYWHAHDESVQNGVITSK  
-----ASTVLQPINPTESYGPQIPITGVSRPSSSDISHTTADPS

FVDNVA TLKILOTDL SHSGOYSCSASNPLGTASSARLTAREPKSPFFD  
KAAANEKTLNVNIQLDDEGYCTAVNAGITKKFFKLV IET --- PYFL  
FFSUIPSGSH TASSLFHIPNNNTGNFFILRH LGRETIWSRG --- RV-

KPVSIDVIAGESADECHVTGQPMRI'W'KDNKEIRPGNYTIT'CVGN  
QOQKLYPIILKRLTDCSATGTPPTILFMKGK--RINESDEVDIIGS  
QNPHRTPVLRHRHRTVRPAIKGANKVSVQVATEYPGMCHTCPSAEL

## FIG. 6F

gi 1017427 emb CAA62189	TPHLRI LKVGKDGSGYTCQATNDVGVKDMCSAQLSVKEPPKFKVAKLEASK
gi 3328186	T--LVIDNPOKEVEGRYTCTAIENKAGRSEKDMWVEVLLPKLSKEMINVE
608	T--VATRAALSVPSSSHSLPKTNVGVIAEESTTVVKKPLLLFKDKQND
gi 1017427 emb CAA62189	VAKGESIQLECKISGSPKIKVSWFRNDSELHESWKYNSMFINSVALLTI
gi 3328186	VQ--AGDEPILLECFIEDTSGVHHI TWSRQFGKQGLDMRAQSSSDKSKLY-I
608	IE----IITTTTKYSGGESNHVIPLEAS-----MTSAPTSVSLG--
gi 1017427 emb CAA62189	NEASAEKSDGYICBAHNGVGDAASCALTIVKAPP-VFTOKPSPVGLKGS
gi 3328186	MQATPEDADSYSCIAVNDAGGAEAVFVTVTPPKIFGDSFSTTEIVADT
608	--KSPVDNS-----GHLSPGPTIQGKDSVETP---LFSPLSTFSPTS
gi 1017427 emb CAA62189	DVILQCEISGTPPEFVWVKDRKQVRNSKKFKITSKHFDNLHINLEAS
gi 3328186	TLEIPCRTEGIPPEISWELDGK-----PILEMP--
608	TKTSKRTTLHQIFVNNQKKEG-----LKNP--
gi 1017427 emb CAA62189	DVGEYHCKATNEVGS DTCSCSVKFEKPPRFVKKLSDTSTLIGDAVELRAI
gi 3328186	--G-----VTK--Q-----CD-----
608	-----YQFG-----



gi|1017427|emb|CAA621891  
gi|3328186  
608

FIER-LEHVEAVIGEPATLQKVDGTPAIRISWYKEHTKLRSAPAYKMQE  
FVRKGGLYEVIENDITMDCGVTSRLPISISWFRGDKPVLXDYSISP  
PEIK-----YQL-KSYSETIEKKRAVSMSPHLSLPEASTHASHWTQ

## FIG. 6I

yi1017427|emb|CAA62189|  
gi13328186  
608

KNWVASLVIKVDHSDVGXYSCADNSVGAVASSAVLVIKARKLPPEFAR  
DG--SHITINKAKLSDGGKYICRASNEAG---TSDIDLILKILVPPKIDK  
KH-----AEKSVFDKKPG--QNPTSKHLP-----YVSLPXTLLKKPRIIG  
\* \* \* \* \*

gi|1017427|emb|CAA62189|  
gi|3328186  
608

K--LKDVHETLGPVAFECRNGSEPLQVSWYKDGVLKKDDANLQTSFVH  
 SNIIGNPLAIVARTIYLECPSGIPQDPVITWTKGMDIN-MTDSRVILAQ  
 G-G-KAASFTVPANSDVFLPCEAVGDPLPIIHWRVSSGXEISQGTQKSRFH

gi|1017427|emb|CAA62189|  
gi|3328186  
608

NVVA--TLQILTDQSHIGQNCASNPGLGTASSAKLILSEHVPVPPFDI  
 NNNE--TFGIENVQVTDQGRYTCTATRGKASHFDSLDLS---PPEFDI  
 VLVLPNGTLLSIQRVSIQDRGQYLCSAFNPGLVDHFVHSVVPY-ARILDR  
 \* \* \* \* \*

gi|1017427|emb|CAA62189|  
gi|3328186  
608

PGKPVSVDLA-LGSGTFFKCHVTGTAPIK-----ITWAKDNREIR---PGGN  
HIGTQPTIKREGDTITLTCPIKLAEDIAQVMDVSWTKDSRALD---GDLT  
HVHXEITVH-FGSTVELKCRVEGMPRPT-----VSWILANQTVVSETAKGS

gi|1017427|emb|CAA62189|  
gi|3328186  
608

YKWTILVENTATILTVLKVGKDAGQYTCYASNIAGKDCSQAQLGVQEPPE  
DNVDI SDDGRKLTISQASLENAGLYTICIALNRAGEASLEFKVEI LSPVVI  
RKVWVTPDG-TLIIYNLSLYDRGFYKCVASNP SGQDSLVLVKIQVITAPPV  
... . \* . \* . \* . \* . \* . \* . \* . \* . \*

gi|1017427|emb|CAA621891  
gi|3328186  
608

**SUBSTITUTE SHEET (RULE 26)**

FIG. 6K

gi|1017427|emb|CAA62189|  
 gi|3328186  
 608

gi|1017427|emb|CAA62189|  
 gi|3328186  
 608

gi|1017427|emb|CAA62189|  
 gi|3328186  
 608

gi|1017427|emb|CAA62189|  
 gi|3328186  
 608

gi|1017427|emb|CAA62189|  
 gi|3328186  
 608

--WISYSENATLOFSRVPANAGKYTCQIRNDAGMOECFATLSVLEPAT  
 ...GITVASDSTFTQINNVSLSDKGVYTCVAENVAGSDNIMYNVDVQ-AP  
 RTKRYTLFHNGTLYFNNGVHAEGBDYICSAQNTLGKDEMKVHLTALTAIP  
 : : : : : \* \* \* \* \* : : : : :  
 IVEKPE5--IKVTTGDTCTLECTVAGTPELSTKWFKDGKELTSDNKYKIS  
 VISNGGT--KQVIEGELAVTECLVEGYPAQVSWLFRNGRVETG--VOGVR  
 RIROSYKTTMRAGETAVALDCEVTGEPKENVFWLLPSNNVISE--SNDRF  
 : : : : : \* \* \* \* \* : : : : :  
 FTKVSGGLKIINVAFSDSGLVSEVQNPVQKDSCTASLOVSDRTVPPSPFT  
 YVTGDRMLTIIEARSLDSGLYICSAATNEAG--SAQOAYTLEVLVSPKILT  
 TTHANRTLSIHKVKPLDSGDYVCVAQNPSG--DDTKYKLDIVSKRPLIN  
 \* \* \* \* \* : : : : :  
 RKLK-----ETNGLSGSSVVMCKVYGSPPISVSWFEGNEGLISSGRKYQT  
 STPG-----VLTPSSGSKFSLPCAVRGYPDPFIISWTINGNDIKDGENGHT  
 GLYANKTVIKATAIRHSKKYDFCRADGIPSPSQVTWIMPNGNIFLAPYFGS  
 : : : : : \* \* \* \* \* : : : : :  
 TLTDNIC-ALTVNMLESDSGDYTCIATNMAGSDECSAPL-----TVREPP  
 IGADG---TLHIEKAEER-HLIYECTAKNDAGADTLEFPVQ---TIVAPK  
 RVTVHPNGTLEMRNIRLSDSADFTCVVRSEGESVLVVQLEVLLEMLRRPT  
 : : : : : \* \* \* \* \* : : : : :  
 : : : : : \* \* \* \* \* : : : : :



FIG. 6L

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

31VQKPDMDVLGTINVTFTSIVKGTTFPSVSWFK-GSELVPGDRCNV5  
ISTSGNRVINGSEGTETVIKCEIE-SESSEFSWSK-NGVPLPSN-NLI  
FRNPFNEKVIQAQKGPVALNCSVDGNPPPEITWILFDGTQFANRPHNSPY  
\* \* \* \* \*

LEDSVAELELEFDVDTSQSGEYTCIVSNEAGKASCTTHLYIKAPAKFKRL  
FSEDYKLKILSTRLSDOGESCTAANKAGNATOKTNLNVGVAPKTIERP  
LMAAGSLILYKATRNKSGKYCAARNKVGYIEKLILLEIGQKPVLLTYE  
\* \* \* \* \*

NDXSIEK-GKPLILEGTTGTPPISVTWK-KNGINVTPSQRCNITTTTKS  
RTOVVKH-GDQVTLWCEASGVFPQAITWY-KDNELLTNGVDETATTKK  
FGWVKSVSGEPLSLHCVSDBGIPKPNKWTTPGGHVIDRPOVDGKYILHN  
\* \* \* \* \*

PILEIPSSSTVEDAGQVNCYIENASGKDCSAQIILILEPPYFVKQLEPVK  
SVIFSSITPSQ-AGVYTCRAENWASTEEDILIVMTPEVVPERNVST  
GTLVIKATTAHQGNVICRAQMSVGQAVISVSVNMVVAYPRIINYLPRNM  
\* \* \* \* \*

5 VGDASASLOCAGTPEIGVSWYKG-DTKLRPTTTYKMHFRNV---A  
N--PRQTVELSCNRNTGTPPEVTSWNR--DSNIAIGNNKYQILG-----T  
LRRTGEAMQLHCVALGIPKPKVKTWETPRHSILSKATARKPHRSEMLHPQG  
\* \* \* \* \*



## FIG. 6N

gi|1017427|emb|CAA62189|  
 gi|3328186  
 608

ETEGNSFKLEGRVAGSQPIITVAWYKNNIEIQTSNCEITFKNN--TIVLQ  
 CYERNOAYSRG-----LTWEYNGVP-MPKNLAGIHEMNGSIVILD  
 -----

gi|1017427|emb|CAA62189|  
 gi|3328186  
 608

VRKAGMNDAGLYTCKVSNDAQSALCTSSIVIKEPKPPVFDQHLTPVTVS  
 TSSLKEGDLELYTCKVNRN-----RHSIPHLTS  
 -----

gi|1017427|emb|CAA62189|  
 gi|3328186  
 608

EGEYVOLSCHVGQSEPIRQWLKAGREIKPSDRCSFASGTAVLELRDV  
 AFEQVP-----EVTID-----KVEV  
 -----

gi|1017427|emb|CAA62189|  
 gi|3328186  
 608

AKADSGDYVCKASNVAGSDTTTSKVTIKDKPAVAPATKKAADVGRLEFFVS  
 NNGDSVILDC-----EVT-----S  
 -----

gi|1017427|emb|CAA62189|  
 gi|3328186  
 608

EPOSIRVVEKTTATFTAKVGGDPIPNVKWTKGKWRQINOGGRVFIIHOKGD  
 DP--LTT-----HVVWTKNDQKMLDDDAIYVLP  
 -----

FIG. 60

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

EAKLEIRUTTKTDSGLYRCVAFNEHGETESWNLOVDERKKOEKTEGDILR  
NNSLVLLNVEKYDEGVKCVASNSIGKAFDDTQLNVY-----EGDFL  
-----

AMLKKTPIIKKGAGEEEIDIMELLKNVDPKYEKVARMYGITDFRGLIQ  
PLTG-----FE-GSG-----IN-IDD-----SSNAGGSSRR-----  
-----

RVVEEHRVKEVHRVIEVFEEVEVEFEKPKAPKGPPEISEKIIPPKQPT  
-----

KVVPKPEPAKVPEVPKIVVEEKRVPEBPVPTKVPVLPPKEVVPE  
-----

KNVPVPPAKKPEAPPKVPPEAPKEVWFEKKVPVPPKKPEVPPTKVPPEVP  
-----

## FIG. 6P

gi|1017427|emb|CAA62189|  
gi|3328186  
608

KVAVPEKKVPEAIPPKPESPPPEVFEEPEESTSPAPPKKPEVPPVRVPEVP  
-----

gi|1017427|emb|CAA62189|  
gi|3328186  
608

KEVVPKKVPAAPPKKPEVTVKVPEAPKEVVPEKKVFPVPPKKPEVPPPT  
-----

gi|1017427|emb|CAA62189|  
gi|3328186  
608

KVPEVPKVAVPEKKVPEAIPPKPESPPPEVFEEPEESTSPAPPKKPEVPPPT  
-----

gi|1017427|emb|CAA62189|  
gi|3328186  
608

EPAAPEQVTVPPKQNPVPEKKGAPAVVAKKPELFPVKVPEVPKEVVPEKKVP  
-----

gi|1017427|emb|CAA62189|  
gi|3328186  
608

LVPVKKPEAPPAKVPEVPKEVVPEKKVAVPKKPEVPPAKVPEVPKKPVLE  
-----

FIG. 6Q

gi|11017427|emb|CAA62189|  
 gi|3328186  
 608

gi|11017427|emb|CAA62189|  
 gi|3328186  
 608

gi|11017427|emb|CAA62189|  
 gi|3328186  
 608

gi|11017427|emb|CAA62189|  
 gi|3328186  
 608

gi|11017427|emb|CAA62189|  
 gi|3328186  
 608

EKDAVPVPERAESPPVEVEPEEIAPEEEIAPEEEKPVAEEEEPEVP  
 -----

PPAVPEEPKCIIPKCKVVIKTPAAPPKPEPEPKVIEKPKLKRPPPP  
 -----

PAPPKEDVKEKTFQLKAI PKCKVPENPQVPEKVELTFLKVPGGCKKVKRL  
 -----

LPERKPEPKEVVLSVLKRKPPEEEKPEPKKLEKVKKPAVPEPPPPK  
 -----

VEEVEVPTVTKRERKIPEPTKVPEIKPAIPLFAPEPKPKPEAEVKTIKP  
 -----

## FIG. 6R

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

gi|1017427|emb|CAA62189|  
gi|3328186  
608

PVEPEPTPIAAPVTVPVGGKAEAKAPKEEAAKPGPIKGVPKKTSPTIE  
-----

AEERRKLRPGSGGKPPDEAPFTYQLKAVELFVKEIKDIIILTESEFVGSS  
-----

AIFFELVSPSTAITTWMKGSNIRESPKHRTIADGKDKRLHIIDVQLSDA  
-----

GEYTCVLRIGNKEKSTAKLVVEELPVRVFKTLEEEVTVVKGQPLYLSCE  
-----

INKERDVWRKDGKIVVEKPGRIVGVIGLMRALTIINDAD  
-----

91/183

405 Human

PROBE: Human 405 EST Probe.

Target: total RNA: Human K562

M K562

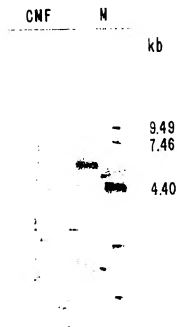


FIG. 7



92/183

405 Human

Probe: Human 405 EST ProbeTarget: Rat Cmf mRNA**FIG. 8**



## FIG. 9B

W	F	L	P	P	V	C	M	A	H	C	I	R	L	A	L	I	R	F	H
TGG	TTC	CTG	CCT	CGG	GTC	TGC	ATG	GCC	CAC	TGC	ATC	CGC	CTG	CGG	CTC	ATC	CGC	TTC	CAC
245					254			263			272			281			290		
V	K	V	A	Q	S	I	E	D	H	H	Q	E	V	I	G	F	C	R	E
GTC	AAG	GTT	GCA	CAG	AGC	ATT	GAG	GAT	CAT	CAT	CAT	CAT	GAA	GTG	ATT	GGC	TTC	TGC	AGA
305					314			323						332			341		350
N	G	F	H	G	L	V	A	Y	D	S	D	Y	A	L	C	N	I	P	Y
AAT	GGT	TTC	CAT	GGC	TTG	GTG	GGG	TAT	GAC	TCT	GAT	TAT	GCA	CTG	TGC	TGC	AAC	ATC	CCC
365					374			383						401			410		
Y	F	S	A	H	A	L	K	L	S	R	N	G	K	S	L	T	T	S	Q
TAC	TTC	AGT	GCC	CAT	GCC	CTG	AAA	CTG	AGC	CGC	ATG	GGG	AGG	AGT	CTC	ACC	ACA	AGC	CAA
425					434			443						461			470		
Y	L	M	H	E	V	A	K	Q	L	D	L	N	P	N	R	F	P	I	F
TAT	CTG	ATG	CAC	GAG	GTT	GCC	AAG	CAA	CTG	GAC	CTG	AAC	CCA	CCA	CGT	TTT	CCT	ATT	TTT
485					494			503						512			521		530

## FIG. 9C

A	A	L	G	N	H	I	L	P	D	E	D	L	A	S	F	H	W	S		
GCT	GCT	CTG	TTA	GGA	AAT	CAC	ATA	CTA	CCT	GAC	GAG	GAT	CTG	GCT	TCC	TTT	CAT	TGG	AGT	
		545		554		554		563		572		572		581				590		
L	L	G	P	E	H	P	L	A	S	L	K	V	R	A	H	Q	L	V	L	
TTA	CTT	GGT	CCA	GAG	CAT	CCG	CTC	TCA	CGC	TCA	CTG	AAG	GTC	CGA	GCT	CAC	CAG	CTG	GTC	TTG
		605		614		614		623		623		632		641				650		
P	P	C	D	V	V	I	K	A	V	A	D	Y	V	R	N	I	H	D	T	
CCA	CCA	TGT	GAT	GTA	GTG	ATC	AAA	GCT	GTA	GCT	GAT	TAT	GTC	CGC	AAC	ATC	CAT	GAC	ACG	
		665		674		674		683		683		692		701				710		
S	D	L	D	A	I	A	K	D	V	F	Q	H	S	Q	S	R	T	D	D	
TCT	GAC	CTG	GAT	GCC	ATA	GCT	AAA	GAT	GTT	TTC	CAG	CAT	TCA	CAG	TCT	AGA	ACA	GAT	GAC	
		725		734		734		743		743		752		761				770		
K	V	I	R	F	K	R	A	V	G	Y	Y	S	A	T	S	K	P	M	P	
AAA	GTC	ATT	CGA	TTT	AAG	AGA	GCA	GTT	GGA	TAT	TAC	TCA	GCA	ACT	AGC	AAG	CCT	ATG	CCA	
		785		794		794		803		803		812		821				821		

## FIG. 9D

F	H	P	P	H	Y	L	P	P	N	P	F	G	M	P	G	M	V	P
TTT	CAC	CCA	CCA	CAT	TAC	TTA	CCC	CGA	CCA	AAT	CCA	TTT	GGA	ATG	CCT	GGG	ATG	GTG
		845			854			863			872			881			890	
P	Y	V	P	P	Q	M	L	N	I	P	Q	T	S	L	Q	A	K	P
CCA	TAT	GTT	CCC	CCT	CAG	ATG	CTC	AAC	ATT	CCA	CAG	ACT	TCT	CTG	CAA	GCA	RAG	CCC
		905			914			923			932			941			950	
V	P	Q	V	P	S	P	G	G	T	P	G	Q	A	P	Y	P	Y	S
GTC	CCA	CAG	GTG	CCC	AGC	CCA	GGG	GGC	ACC	CCA	GGC	CAG	GCT	CCA	TAC	CCA	TAC	AGC
		965			974			983			992			1001			1010	
P	E	P	A	L	A	L	D	S	S	G	K	N	L	T	E	Q	N	S
CCT	GAG	CCT	GCC	CTC	GCC	CTG	GAC	TCT	AGC	GGG	AAG	AAC	CTG	ACA	GAG	CAG	AAC	AGC
		1025			1034			1043			1052			1061			1070	
S	N	I	P	H	E	G	K	H	T	P	L	Y	E	R	S	S	P	I
AGC	RAC	ATT	CCT	CAC	GAG	GGG	AAG	CAC	ACA	CCG	CTG	TAT	GAG	CGC	TCC	TCG	CCC	ATC
		1085			1094			1103			1112			1121			1130	

## FIG. 9E

L A Q S G S P N H V D S A Y F P G S S T  
 CTG GCC CAG AGC GGC AGC CCC AAT CAC GTG GAT TCC GCC TAC TTC CCT GGC TCG TCT ACC  
 1145 1154 1163 1172 1181 1190  
  
 S S S S D N D E G S G G A T N H I S G N  
 TCG TCC TCA TCA GAC AAC GAT GAG GGC AGT GGA GGC GCA ACC AAC CAC ATC AGC GGC AAC  
 1205 1214 1223 1232 1241 1250  
  
 K I G W E K T G S H A E P L A R G D P G  
 AAG ATT GGC TGG GAG AAG ACA GGA AGC CAT GCA GAG CCA CTG GCA CGG GGA GAC CCA GGA  
 1265 1274 1283 1292 1301 1310  
  
 D Q V K V E G S S T A S S G S Q L A E G  
 GAC CAA GTA AAG GTA GAA GGT TCA TCC ACT GCC TCT TCG GGA AGC CAG CTA GGT GAA GGC  
 1325 1334 1343 1352 1361 1370  
  
 K G S H M G T V Q P I P C L L S M P T R  
 AAA GGG AGC CAC ATG GGC ACT GTT CAG CCA ATC CCG TGC CTC CTG TCA ATG CCC ACC AGG  
 1385 1394 1403 1412 1421 1430

## FIG. 9F

N H M D I T P P L P P P V A P E V L R V  
 AAC CAC ATG GAT ATC ACC ACA CCT CCC TTG CCC GTT GCA CCT GAG GTG CTG AGA GTG  
 1445 1454 1463 1472 1481 1490

A E H R H K K G L M Y P Y I F H V L T K  
 GCT GAG CAC AGA CAC AAG AAG GGA CTG ATG TAC CCC TAC ATC TTC CAC GTC CTG ACA AAG  
 1505 1514 1523 1532 1541 1550

G E I K I A V S I E D E A N K D L P P A  
 GGT GAA ATC AAA ATT GCT GTT TCT ATT GAA GAT GAA GCC AAC AAA GAC CTG CCT CCG GCC  
 1565 1574 1583 1592 1601 1610

A L L Y R P V R Q Y Y G V L F S L A E  
 GCC CTG CTC TAT AGG CCA GTT CGT CAG TAT GTT TAC GGA GTC CIG TTT AGT TTG GCA GAA  
 1625 1634 1643 1652 1661 1670

S R K K T E R L A F R K N R L P P E F S  
 AGC AGA AAG AAA ACT GAG AGA CTT GCT TTT AGA AAG AAC AGA CTT CCA CCA GAA TTC TCG  
 1685 1694 1703 1712 1721 1730

## FIG. 9G

P L I I K E W A A Y K G K S P Q T P E L  
 CCA CTG ATC ATT AAG GAA TGG GCA GCT TAC AAG GGG AAG TCT CCT CAG ACC CCG GAA CTA  
 1745 1754 1763 1772 1781 1790

V E A L A A F R E W T C P N L K R L W L G  
 GTG GAA GCA CTG GCC TTC CGG GAG TGG ACC TGC CCC AAC CTG AAG AGG CTT TGG CTG GGC  
 1805 1814 1823 1832 1841 1850

K A V E D K N R R M R A F L A C M R S D  
 AAG GCG GTG GAG GAC AAG AAC CGC CGG ATG AGG GCC TTC CTG GCC TGC ATG AGG TCT GAT  
 1865 1874 1883 1892 1901 1910

T P A M L N P A N V P T H L M V L C C V  
 ACC CCA GCC ATG CTC AAC CCT GCC AAT GTG CCC ACT CAC CTC ATG GTG CTC TGC TGT GTC  
 1925 1934 1943 1952 1961 1970

L R Y M V Q W P G A R I L R R Q E L D A  
 TTA CGG TAT ATG GTG CAG TGG CCC GGA GCT CGC ATT CTG CGG CGT CAG GAG CTG GAT GCC  
 1985 1994 2003 2012 2021 2030



## FIG. 9H

F L A Q A L S P K L Y E P D Q L Q E L K	
TTC CTG GCT CAG GCA TTG TCT CCC AAA CTC TAT GAG CCG GAC CAG CTG CAG GAA CTC AAG	2045 2054 2063 2072 2081 2090
I E N L D P R G I Q L S A L F M S G V D	
ATT GAG AAC TTA GAC CCC CGA GGA ATT CAG CTG TCA GCT CTC TTC ATG AGC GGA GTA GAC	2105 2114 2123 2132 2141 2150
M A L F A N D A C G Q P I P W E H C C P	
ATG GCA CTG TTT GCC AAT GAC GCG TGT GGG CAG CCC ATA CCC TGG GAA CAC TGT TGT CCT	2165 2174 2183 2192 2201 2210
W M Y F D G K L F Q S K L L K A S R E K	
TGG ATG TAT TTC TAT GCG AAG CTC TTC CAG TCC AAA CTT CTT AAA GCC AGC CGG GAA AAG	2225 2234 2243 2252 2261 2270
T P L I D L C D G Q A E Q A A K V E K M	
ACC CCA CTC ATC GAC CTC TGT GAT GGT CAG GCT GAG CAG GCT GCC AAG GTT GAA AAG ATG	2285 2294 2303 2312 2321 2330

## FIG. 9I

R Q S I L E G L N F S R Q N H P L P F P	
CGC CAG AGT ATC CTG GAG GGG CTC AAC TTC TCC CGG CAA AAC CAC CCA CTC CCC TTC CCA	2390
	2381
	2372
	2363
	2354
	2345
P P A L P F Y P A S V Y P R H F G P V	
CCA CCT CCT GCT CTG CCC TTC TAC CCA GCT TCT GTA TAT CCT CGG CAC TTT GGG CCG GTC	2450
	2441
	2432
	2423
	2414
	2405
P P S Q G R G R G F A G V C G F G G H Y	
CCA CCC TCA CAG GGC AGG GGC AGG GGT TTT GCA GGA GTC TGT GGC TTT GGA GGC CAT TAT	2510
	2501
	2492
	2483
	2474
	2465
G E T V A T G P Y R A F R V T A A S G H	
GGG GAA ACT GTA GCA ACA GGC CCT TAC GCT GCC TTC CGG GTG ACA GCA GCA TCG GGA CAC	2570
	2561
	2552
	2543
	2534
	2525
C G A F S G S D S N R T S K S Q G G V Q	
TGT GGA GCC TTC TCA GGC AGT GAC AGC AAC AGG ACT AGC AAG TCC CAG GGC GGA GTC CAG	2630
	2621
	2612
	2603
	2594
	2585

## FIG. 9J

P I P S O G G K L E I A G T V V G H W A  
 CCT ATT CCT TCT CAG GGA GGG AAG CTA GAA ATA GCT GGC ACG GTG GTC GGC CAC TGG GCT  
 2645 2654 2663 2672 2681 2690

G S R R G R G G G G C G T G G G C G A G G G C C C T T C C T C T G C A G G T G T G T C T G T G G A  
 2705 2714 2723 2732 2741 2750

G P A R G G R P R G V I S T P V I R T F G  
 GGA CCA GCA AGA GGG CGT CCA AGG GGT GTT ATT TCC ACT CCT GTG ATT AGA ACA TTT GGA  
 2765 2774 2783 2792 2801 2810

R G G R Y Y G R G Y K S Q G A I Q G R P  
 AGA GGT GGA AGG TAC TAC GGC AGA GGC TAT AAA AGC CAG GGA CCA ATT CAG GGC AGA CCT  
 2825 2834 2843 2852 2861 2870

P Y A A S A E E V A K E L K S K S G E S  
 CCT TAT GCT GCT TCA GCA GAA GAA GTG GCC AAA CTG AAA TCG AAA TCT GGG GAA TCC  
 2885 2894 2903 2912 2921 2930

## FIG. 9K

K S S A V S S E G S L A E N G V M A E E  
 AAG TCC TCT GCT TCA GAA GGG TCC CTG GCT GAA AAC GGA GTG ATG GCC GAG GAG  
 2945 2954 2963 2972 2981 2990  
  
 K P V P Q L N G S T A D T R A P S H S E  
 AAG CCA GTG CCC CAA CTT AAT GGG AGC AGC GCT GAC ACC AGG GCC CCC AGC CAC TCT GAA  
 3005 3014 3023 3032 3041 3050  
  
 S A L N N D S K T C N T N P H L N A L S  
 AGT GCC TTG AAC AAT GAC TCT AAA ACG TGC AAT ACA AAT CCT CAC TTA AAT GCA CTA AGT  
 3065 3074 3083 3092 3101 3110  
  
 T D S A C R R E A A L E A A V L N K E E  
 ACA GAC AGC GCC TGC CGC AGG GAG GCT GCT CTG GAG GCA GCT GTC TTA AAT AAA GAA GAG  
 3125 3134 3143 3152 3161 3170  
  
 TGA ACT TAT TTT TAT AGA GGG TGA AGG ATG CTG GAA GGG TAA GGA TTC AGG AAT ATC TGG  
 3185 3194 3203 3212 3221 3230

*FIG. 9L*

AGA GAA AGA GAG CCT GCA GTT ATG TAC ATT TTG TCC TTT CCG TAA GAG AAA ATG AGG ACT  
 3245 3254 3263 3272 3281 3290

GTG GAA ATT CAG ATC CCT CTT TGA TAT CAG AGA TTT AAA CAA CAC ATT CTT TTT TTA GTT  
 3305 3314 3323 3332 3341 3350

TTA ACC AGT TGT AGT CAA AAT GCT ACA ATA AAA CAA AAC GAG AGA GAG AAA ATG AAG AGC  
 3365 3374 3383 3392 3401 3410

ATT TGA CTC CCG CAC TTA AAA ATG AAG CAC ACA AAG TTT AAA CTG GTT ACG ACG ACA AAA  
 3425 3434 3443 3452 3461 3470

*FIG. 9M*

GCC TAC AGT TGT GTT TCT TGA ACT ATA AAG AAA ACA AAT TTT GGC AGT GTT TAA GTA TAT  
 3485 3494 3503 3512 3521 3530

ATA GCT TAA AAT ATA ATT TTT AGC ATT TGG CAC CAT ATG TAT GCC ATT ATA TTT GAT TTT  
 3545 3554 3563 3572 3581 3590

GCA TTA CTG TTT CAC AAT GAA GCT TTG TTT TAA GGC TTT GAT TTA TGA TGA AAG AAA  
 3605 3614 3623 3632 3641 3650

TAA GGC ACA ACC ACA GTT TTT CTT TCT TAC TTA AAT TTC ATC ACT GTT GAT GTG GTT CTT  
 3665 3674 3683 3692 3701 3710

TTG TGA

*FIG. 10A*

Length of r84695-#85 like R405: 4680 bp; Listed from: 2 to: 4680;  
Translated from: 152 to: 3505 (Entire region);  
Genetic Code used: Universal;

## Frame 2

CCG	CGG	CGG	CCA	TGA	GCG	CGC	CCC	CGA	CCC	GCC	CCA	GTC	CCC	CCT	AGA	GGC	CGC	55
	10				19			28				37			46			
CGC	CCC	CGC	CCG	CCA	GCC	CGC	CCG	CGC	GCC	ACG	GCC	CCA	CCC	CCG	GCC	CCG	CGC	CCC
	64			73			82			91		100			109			
CCC	GCC	CGC	ACC	CGC	GCC	CGC	CCC	GCC	GCC	GCC	CCC	ATG	GCC	GTG	CAG	GGC	TTC	CAG
	124			133			142			151		160			169			
Y	I	E	K	H	C	P	S	A	V	V	E	L	Q	K	L	A	R	
TAC	ATC	GAG	AAG	CAC	TGC	CCG	AGC	GCC	GTG	GTG	GAG	CTG	CAG	AAG	CTG	GCC	CGG	
		184			193			202		211		220			229			

## FIG. 10B

G	S	L	V	G	G	G	R	Q	R	P	P	Q	T	P	L	R	L	L	V
GGC	AGC	CTG	GTG	GAC	GAC	GGG	GGG	CAG	CGG	CCC	CGC	CAG	ACC	CGG	CTG	CGC	CTG	CTG	GTG
		244		253		253		262		271		271		280					289
D	A	D	N	C	L	H	R	L	Y	G	G	F	Y	T	D	W	V	S	G
GAC	GCC	GAC	AAC	TGC	CTG	CAC	CGC	CTC	TAC	GCC	GCC	TTC	TAC	ACC	GAC	TGG	GTC	AGC	GGC
		304		313		313		322		331		331		340					349
G	Q	W	N	H	M	L	G	Y	L	A	A	L	A	K	A	C	F	G	G
GGC	CAG	TGG	TAC	CAC	ATG	CTT	GGC	TAC	CTG	GGC	GGC	CTG	GCC	AAG	GCC	TGC	TTC	GGC	GGC
		364		373		373		382		391		391		400					409
N	I	E	L	F	V	F	F	N	G	A	L	E	K	A	R	L	H	E	W
AAC	ATC	GAG	CTC	TTC	TTC	TTC	TTC	AAC	GCC	GCG	CTC	GAG	AAG	GCC	CGG	CTG	CAC	GAG	TGG
		424		433		433		442		451		451		460					469
V	K	R	Q	G	N	E	R	Q	T	A	Q	Q	I	V	S	H	V	Q	N
GTC	AAG	CGG	CAG	GAC	GAG	GAG	CGC	CAG	ACG	GCA	CAG	CAG	ATC	GTC	AGC	CAT	GTC	CAG	AAC
		484		493		493		502		511		511		520					529



## FIG.10C

K	G	T	P	P	P	K	V	W	F	L	P	P	V	C	M	A	H	C	I
AAG	GOC	ACC	CCG	CCG	CCA	AAG	GTC	TGG	TTC	CTG	CCG	CCC	GTC	TGC	ATG	GCC	CAC	TGC	ATC
		544				553		562			571				580			589	
R	L	A	L	I	R	F	H	V	K	V	A	Q	S	I	E	D	H	H	Q
CGC	CTG	GCG	CTC	ATC	CGC	TTC	CAC	GTC	AAG	GTT	GCA	CAG	AGC	ATT	GAG	GAT	CAC	CAT	CAG
		604			613			622			631				640			649	
E	V	I	G	F	C	R	E	N	G	F	H	G	L	V	A	Y	D	S	D
GAA	GTG	ATT	GGT	TTC	TGC	AGA	GAG	AAT	GGT	TTC	CAT	GGC	TTC	GTC	GAT	TAT	GAT	TCT	GAT
		664			673			682			691				700			709	
Y	A	L	C	N	I	P	Y	Y	F	S	A	H	A	L	K	L	S	R	N
TAT	GCA	CTG	TGC	AAC	ATC	CCC	TAC	TAT	TTC	AGT	GCC	CAT	GCC	CTA	AAA	CTG	AGC	CGG	AAC
		724			733			742			751				760			769	
G	K	S	L	T	T	S	Q	Y	L	M	H	E	V	A	K	Q	L	D	L
GGG	AAA	AGT	CTC	ACC	ACA	AGC	CAA	TAT	CTG	ATG	CAT	GAA	GTT	GCC	AAG	CAA	CTG	GAC	CTG
		784			793			802			811				820			829	

## FIG. 10D

N	P	N	R	F	P	I	F	A	A	L	L	G	N	H	I	L	P	D	E
AAT	CCA	AAT	CGT	TTT	CCT	ATT	TTT	GCT	GCT	CTC	TTA	GGA	AAT	CAC	ATT	CTG	CCT	GAT	GAA
844			853				862				871			880		889			
D	L	A	S	F	H	W	S	L	L	G	P	E	H	P	L	A	S	L	K
GAT	CTG	GCT	TCC	TTT	CAT	TGG	AGT	TTA	CTT	GGT	CCA	GAA	CAT	CCA	CTA	GCC	TCA	CTA	AAG
904			913				922				931			940		949			
V	R	A	H	Q	L	V	L	P	P	C	D	V	V	I	K	A	V	A	D
GTC	CGG	GCC	CAC	CAG	CTG	GTC	TTG	CCA	CCT	TGC	GAC	GTA	GTG	ATC	AAA	GCC	GTT	GCT	GAC
	964				973			982			991			1000		1009			
Y	V	R	N	I	Q	D	T	S	D	L	D	A	I	A	K	D	V	F	Q
TAT	GTA	CGC	AAC	ATT	CAG	GAC	ACC	TCT	GAC	TTG	GAT	GCC	ATA	GCT	AAA	GAT	GTT	TTC	CAG
1024					1033			1042			1051			1060		1069			
H	S	Q	S	R	T	D	D	K	V	I	R	F	K	R	A	I	G	Y	Y
CAT	TCA	CAG	TCT	AGA	ACT	GAC	GAC	AAA	GTT	ATT	GCA	TTT	AAG	AGA	GCA	ATT	GGA	TAT	TAT
1084					1093			1102			1111			1120		1129			

## FIG.10E

S A T S K P M S F H P P H Y L A A R P G  
 TCA GCG ACT AGT AAG CCT ATG TCA TTT CAT CCA CCA CAT TAC TTA GCA GCC AGA CCC GGT  
 1144 1153 1162 1171 1180 1189  
  
 P F G M P G M V P P H V P P Q M L N I P  
 CCG TTT GGA ATG CCT GGG ATG GTG CCG CCG CAT GTT CCT CCT CAG ATG CTC AAC ATT CCG  
 1204 1213 1222 1231 1240 1249  
  
 Q T S L Q A K P V A P Q V P S P G G A P  
 CAG ACC TCT CTG CAA GCA AAG CCC GTG GCC CCA CAG GTG CCC AGC CCA GGG GGC GCC CCG  
 1264 1273 1282 1291 1300 1309  
  
 G Q G P Y P Y S L S E P A P L T L D T S  
 GGC CAG GGT CCA TAC CCG TAC AGC CTC TCT GAG CCA GCA CCT CTC ACT TTG GAC ACG AGC  
 1324 1333 1342 1351 1360 1369  
  
 G K N L T E Q N S Y S N I P H E G K H T  
 GGG AAG AAT CTG ACG GAG CAG AAC AGC TAC AGC AAC ATT CCT CAC GAA GGG AAG CAC ACG  
 1384 1393 1402 1411 1420 1429

## FIG. 10F

P L Y E R S S P I N R A Q S G S P N H V  
 CCG CTG TAT GAG CGG TCC TCG CAC ATC AAC CGG GCC AGC AGC CCC AAC CAC GTG  
 1444 1453 1462 1471 1480 1489  
  
 D S A Y F P G S S T S S S D N D E G S  
 GAT TCC GCC TAC TTC CCT GGC TCT TCT ACA TCG TCA TCT TCC GAC AAC GAC GAG GGC AGC  
 1504 1513 1522 1531 1540 1549  
  
 G G A T N H I S G N K I G W E K T G S H  
 GGA GGG GCG ACA AAC CAT ATC AGC GGG AAC AAG ATT GGC TGG GAG AAG ACG GGA AGC CAC  
 1564 1573 1582 1591 1600 1609  
  
 S E P Q A R G D P G D Q T K A E G S S T  
 TCA GAG CCT CAG GCA CGA GGA GAC CCA CCA AAG GAC GCA AAG GGC TCG TCC ACT  
 1624 1633 1642 1651 1660 1669  
  
 A S S G S Q L A E G K G S Q M G T V Q P  
 GCC TCT TCA GGA AGC CAA CTA GCC GAA GGC AAG GGA AGC CAG ATG GGC ACT GTG CAG CCA  
 1684 1693 1702 1711 1720 1729

## FIG. 10G

I P C L L S M P T R N H M D I T P P L  
 ATC CCG TGC CTC CTG TCG ATG CCC ACC AGG AAC CAC ATG GAC ATC ACC ACA CCT CCC CTG  
 1744 1753 1762 1771 1780 1789

P P V A P E V L R V A E H R H K K G L M  
 CCC CCC GTC GCA CCT GAG CTG CTG AGA CTG GCC GAG CAC AGG CAC AAG AAG GAG CTG ATG  
 1804 1813 1822 1831 1840 1849

Y P Y I F H V L T K G E I K I A V S I E  
 TAC CCC TAC ATC TTC CAT GTC CTG ACG AAG GGT GAA ATC AAA ATT GCT GTT TCT ATT GAA  
 1864 1873 1882 1891 1900 1909

D E A N K D L P P A A L L Y R P V R Q Y  
 GAT GAA GCC AAC AAG GAC CTG CCT CCG GCC GCT CTG CTC TAT AGG CCA GTT CGT CAG TAT  
 1924 1933 1942 1951 1960 1969

V Y G V L F S L A E S R K K T E R L A F  
 GTT TAC GGA GTC TTT AGT TTG GCA GAA AGC AGA AAG AAA ACT GAG AGA CTT GCT TTT  
 1984 1993 2002 2011 2020 2029

## FIG. 10H

R	K	N	R	L	P	P	E	F	S	P	V	I	I	K	E	W	A	A	Y
AGA	AG	AG	AG	AGA	CTT	CCA	CCA	GAA	TTT	TCA	CCA	CTG	ATC	ATT	AAA	GAA	TGG	GCA	GCT
2044					2053			2062			2071			2080			2089		
K	G	K	S	P	Q	T	P	E	L	V	E	A	L	A	F	R	E	W	T
AAA	GGA	AG	TCT	CCT	CAA	ACC	CCG	GAA	CTG	GTT	GAA	GCT	CTT	GCC	TTT	AGG	GAG	TGG	ACC
2104				2113				2122			2131			2140			2149		
C	P	N	L	K	R	L	W	L	G	K	A	V	E	D	K	N	R	R	M
TGC	CCC	AAC	CTG	AAG	AGG	CTG	TGG	TTG	GGT	AAG	CGG	GTA	GAG	GAC	AAG	AAC	COC	AGG	ATG
2164				2173				2182			2191			2200			2209		
R	A	F	L	A	C	M	R	S	D	T	P	A	M	L	N	P	A	N	V
TAGG	GCC	TTC	CTG	GCC	TGC	ATG	AGG	TCG	GAC	ACC	CCA	GCC	ATG	CTC	AAC	CCT	GCC	AAC	GTG
2224				2233				2242			2251			2260			2269		
P	T	H	L	M	V	L	C	C	V	L	R	Y	M	V	Q	W	P	G	A
CCC	ACT	CAC	CTC	ATG	GTG	CTC	TGC	TGC	GTC	TTA	CGG	TAC	ATG	GTG	CAG	TGG	CCG	GGA	GCA
2284				2293				2302			2311			2320			2329		

## FIG. 101

C	I	L	R	R	Q	E	L	D	A	F	L	A	Q	A	L	S	P	K	L
TGC	ATC	CTT	CGG	CGT	CAG	GAG	GAG	CVA	GAT	GCC	TTC	CTG	GCT	CAG	GCG	CTG	TCC	CCC	AAA
		2344			2353				2362		2371			2380				2389	
Y	E	P	D	Q	L	Q	E	L	K	I	E	N	L	D	P	R	G	I	Q
TAC	GAG	CCT	GAT	CAG	CTC	CAG	GAG	CTC	AAG	ATT	GAC	AAC	CTA	GAT	CCC	CGA	GGA	ATT	CAG
		2404			2413				2422		2431			2440				2449	
L	S	A	L	F	M	S	G	V	D	M	A	L	F	A	N	D	A	C	G
CTA	TCA	GCT	CTC	TTC	ATG	AGT	GGA	GTA	GAC	ATG	GCC	TTG	TTT	GCA	AAT	GAT	GCA	TGC	GGA
		2464			2473				2482		2491			2500				2509	
Q	P	I	P	W	E	H	C	C	P	W	M	Y	F	D	G	K	L	F	Q
CCAG	CCA	ATC	CCC	TGG	GAA	CAC	TGT	TGT	CCT	TGG	ATG	TAT	TTT	GAT	GGG	AAG	CTC	TTC	CAA
		2524			2533				2542		2551			2560				2569	
S	K	L	L	K	A	S	R	E	K	T	P	L	I	D	L	C	D	G	Q
CTCC	AAA	CTC	CTC	AAA	GCC	AGC	CGG	GAA	AAG	ACC	CCA	CTC	ATT	GAC	CTC	TGT	GAT	GGT	CAG
	2584				2593				2602		2611			2620				2629	

## FIG. 10J

A	D	Q	A	A	K	V	E	K	M	R	Q	S	V	L	E	G	L	S	F
GCT	GAT	CAG	GCT	GCC	AAG	GTA	GAG	AAG	ATG	CGC	CAG	AGC	CTC	CTC	GAG	GGG	CTC	AGC	TTC
		2644			2653				2662			2671			2680			2689	
S	R	Q	S	H	T	L	P	F	P	P	P	P	A	L	P	F	Y	P	A
TCC	AGG	CAG	ACC	CAC	ACG	CTC	CCT	CTT	CCG	CCG	CCA	CCT	GCC	CTG	CCC	TTC	TAC	CCT	GCC
		2704			2713				2722			2731			2740			2749	
S	A	Y	P	R	H	F	G	P	V	P	P	S	Q	G	R	G	R	G	F
TCT	CGC	TAC	CCC	CGC	CAC	TTT	GGG	CCT	GTC	CCA	CCC	TCT	CAG	GGC	AGG	GGC	AGA	GGC	TTT
		2764			2773				2782			2791			2800			2809	
A	G	V	C	G	F	G	G	P	Y	G	E	T	V	A	T	G	P	Y	R
GCA	GGC	GTC	TGT	GGC	TTT	GGA	GGC	CCC	TAT	GGG	GAA	ACG	GTA	GCA	ACA	GGC	CCT	TAC	CGT
		2824			2833				2842			2851			2860			2869	
A	F	R	V	A	A	A	S	G	H	C	G	A	F	S	G	S	D	S	S
GCC	TTT	CGT	GTC	GGC	GCA	GCA	TGC	GGA	CAC	TGC	GGA	GCC	TTT	TCA	GGC	AGT	GAC	AGC	AGC
		2884			2893				2902			2911			2920			2929	



116/183

## FIG. 10K

R T S K S Q G G V Q P I P S Q G G K L E	
AGG ACT AGC AAG TCC CAG GGC GGA GTC CAA CCT ATA CCT TCT CAG GGA GGC AAA CTA GAA	2944 2953 2962 2971 2980 2989
I A G T V V G H W A G S R R G R G G R G	
ATA GCT GGC ACT GTG GTT GGC CAT TGG GCT GGG AGC AGG CGG GGC CGT GGG GGC CGG GGG	3004 3013 3022 3031 3040 3049
P F P L Q V V S V G G G P A R G R P R G V	
CCT TTC CCC CTG CAG GTG GTT TCT GTC GGA GGA CCA GCT AGA GGG CGT CCA AGA GGA GTT	3064 3073 3082 3091 3100 3109
I S T P V I R T F G R G R Y Y G R G Y	
ATT TCC ACC CCA GTG ATT AGA ACA TTT GGA AGA GGT GGA AGG TAC TAT GGC AGA GGT TAC	3124 3133 3142 3151 3160 3169
K N Q A A I Q G R P P Y A A S A E E V A	
AAA AAC CAG GCA ATT CAG GGC AGA CCT CCT TAT GCT GCT TCA GCA GAA GAA GTG GCC	3184 3193 3202 3211 3220 3229

## FIG. 10L

K E L K S K S G E S K S S A M S S D G S  
 AAA GAA CTT AAG TCA AAA TCT GGG GAA TCG AAG TCC TCT GCT ATG TCT TCA GAC GGG TCC  
 3244 3253 3262 3271 3280 3289

L A E N G V M A E E K P A P Q M N G S T  
 CTG GCT GAA AAC GGA CTG ATG GCC GAG GAG AAG CCG GCT CCG CAG ATG AAC GGG AGC ACG  
 3304 3313 3322 3331 3340 3349

G D A R A P S H S E S A L N N D S K T C  
 GGT GAC GCC AGG GCC CCC AGC CAC TCT GAA AGT GCC TTG AAT AAT GAC TCT AAA ACG TGC  
 3364 3373 3382 3391 3400 3409

N T N P H L N A L S T D S A C R R E A A  
 AAT ACA AAT CCT CAT TTA AAT GCA CTA AGT ACA GAC AGC GCT TGC CGC AGA GAA GCT GCT  
 3424 3433 3442 3451 3460 3469

L E A A V L N K E E  
 CTG GAG GCA GCT CTC TTA AAT AAA GAG TAA ACT TAT TTT TTA TAG AGG GTG AAG GAT  
 3484 3493 3502 3511 3520 3529

118/183

*FIG. 10M*

GCT GGA AGG GTA AGG ATT TRG GAA TAT CTG GAG AGA AAG AGA GCC TGC AGT TAT GTA CAT  
3544 3553 3562 3571 3580 3589

TTT GTC CTT TCC GTA AGA GAA AAA TGA GGA CTT TGG AAA TTC AGA TCC CTC TTT GAT ATC  
3604 3613 3622 3631 3640 3649

AGA GAT TTA AAC AAC ACA TTT TTA ACC AGT TGT AGT CAA AAT GCT ACA ATA AAA  
3664 3673 3682 3691 3700 3709

CAA AAA AGA GAA AGA AAA TGA AGA GCA TTT GAC TCC CGC ACT TRA AAT GAA GTA CAC ATA  
3724 3733 3742 3751 3760 3769

AAG TTT AAA CTG GTT ATG ACA AAA GCC TAT AGT TGT GTT TCT TGA ACT ATA AAG AAA ACA  
3784 3793 3802 3811 3820 3829

*FIG. 10N*

AAT TTT GGC AGT CTT TAA GTA TAT ATA GCT TAA AAT ATA ATT TTT AGC ATT TGG CAC CAT  
3844 3853 3862 3871 3880 3889

ATG TAT GCC ATT ATA TTT GAT TTT GCA TTA CTG TTT CAC AAT GAA GCT TTC TTT AAG GCT  
3904 3913 3922 3931 3940 3949

TTG ATT TTT ATG ATT ATG AAA GAA ATA AGG CAC AAC CAC AGT TTT TCT TTC TTA AAT TTC  
3964 3973 3982 3991 4000 4009

ATC ACT GTT GAT GTG GTT CTT TTG TGT TAA AAA AAA AAA AAA CTA TCA AAA CTA AAA  
4024 4033 4042 4051 4060 4069

AAT TAT AGA GTA ATA TTG CCG TTC TGC TGA TTT TAA ATA TAC AAT ACA TCA TAC ATA CTT  
4084 4093 4102 4111 4120 4129

## FIG. 10 0

TAC	AAG	CAA	GTT	AAA	TGG	AGA	TAA	AGT	TGA	AAT	CAT	AGA	TGC	AAA	TGA	CCT	TTC	AAA	4144	4153	4162	4171	4180	4189
ATC	AAC	ACA	ATG	TGT	TCT	GAA	ACT	TTT	GTG	ACT	AAT	ACC	ATG	CAT	CTG	TGA	TCA	ATG	4204	4213	4222	4231	4240	4249
TAT	GTG	GTT	TTG	AAT	CGG	ATG	TAG	ACC	ATT	AGT	ACT	ACT	ACT	TGA	GCT	AAA	CTT	CTG	4264	4273	4282	4291	4300	4309
GGT	TCA	TAA	TTT	TTA	AAG	TGT	GTA	GTT	AAT	ATG	CAT	GTT	ATC	GTG	CTT	TCT	TCC	ATT	4324	4333	4342	4351	4360	4369
AAC	AGT	ATG	TGC	CCA	TTT	GCA	AAA	CAA	AAA	TGC	TAA	TAA	TCA	GTA	ATA	GTC	CTA	TAA	4384	4393	4402	4411	4420	4429

*FIG. 10P*

ATG TTA ACT' CTG TTT AGT CAT TGA CTG ATC TTG CTC TAA CCT TAA AAT TTT GTG ATT ATT  
 4444 4453 4462 4471 4480 4489

GAC CTC TGT TGC ATT TAT TCT TAA GCC CCC CAA AAA TTA TCT AGC CGT TTC GAA TAT CAA  
 4504 4513 4522 4531 4540 4549

CAT TAC CCT GGT GTA TTC ACT GCT GTA TGC ATT ATT GTT CTT TGT TGC TGT TTT ATG CCT  
 4564 4573 4582 4591 4600 4609

TCA TAT TAG CAA ATA TGA AAT TCT GTG AAA AAA AAC CCT TTG ATC TTA AAA AAA AAA AAA  
 4624 4633 4642 4651 4660 4669

AAA AA



## FIG. IIB

405: CLUSTAL X (1.64b) Multiple Sequence Alignment

```

405_no up
gi|3005744
405_hum
gi|1136426|gml|PID|d1012166

-----
HALKLSRNGKSLTTSQYILMHEVAKQOLDINPNRFFIFAALLGNHILPDEDL
-----
HALKLSRNGKSLTTSQYILMHEVAKQOLDINPNRFFIFAALLGNHILPDEDL
-----
HALKLSRNGKSLTTSQYILMHEVAKQOLDINPNRFFIFAALLGNHILPDEDL
-----

405_no up
gi|3005744
405_hum
gi|1136426|gml|PID|d1012166

-----
ASFHWSLLGPEHPLASLKVRAHQVLVLPDQVVIKAVADYVRNIHDTSDLD
-----
ASFHWSLLGPEHPLASLKVRAHQVLVLPDQVVIKAVADYVRNIQDTSDDL
-----
ASFHWSLLGPEHPLASLKVRAHQVLVLPDQVVIKAVADYVRNIQDTSDDL
-----

405_no up
gi|3005744
405_hum
gi|1136426|gml|PID|d1012166

-----
AIAKDVTFQHSQSRRTDDKVIREFKRAVGYSATSKPMFPHPHYLPRLPNPF
-----
AIAKDVTFQHSQSRRTDDKVIREFKRAIGYSATSKPMSFHPHYLAARPGPF
-----
AIAKDVTFQHSQSRRTDDKVIREFKRAIGYSATSKPMSFHPHYLAARPGPF
-----

405_no up
gi|3005744
405_hum
gi|1136426|gml|PID|d1012166

-----
GNPGMVPPYVPPQMLNIPTSLQAKVAPQVPSGGTTPGQAPYFYSILPEP
-----
GNPGMVPPYVPPQMLNIPTSLQAKVAPQVPSGGAPGGQGPYFYSILSEP
-----
GNPGMVPPYVPPQMLNIPTSLQAKVAPQVPSGGAPGGQGPYFYSILSEP
-----

```



## FIG. 11C

405_no_up	A-LALDSGKILTEQNSYSNI	PHEGKHTPLYERSSPINLAQSGSPNHVDS
gi 3005744	APLTLDTSKNL	TEQNSYSNI PHEGKHTPLYERSSPINRAQSGSPNHVDS
405_hum	APLTLDTSKNL	TEQNSYSNI PHEGKHTPLYERSSPINPAQSGSPNHVDS
gi 1136426 gnl PID d1012166		
405_no_up	AYFPGSSTSSSDNDESGGATNHI	SGNKIGWEKTGSHAEPFARQDQ
gi 3005744	AYFPGSSTSSSDNDESGGATNHI	SGNKIGWEKTGSHAEPFARQDQ
405_hum	AYFPGSSTSSSDNDESGGATNHI	SGNKIGWEKTGSHAEPFARQDQ
gi 1136426 gnl PID d1012166		
405_no_up	VKVGSGSTASSGSQLAEGKGSHMGTVQPI	PCLLSMPTRNHMDITTPPLPP
gi 3005744	TKAEGSSTASSGSQLAEGKGSQMGTVQPI	PCLLSMPTRNHMDITTPPLPP
405_hum	TKAEGSSTASSGSQLAEGKGSQMGTVQPI	PCLLSMPTRNHMDITTPPLPP
gi 1136426 gnl PID d1012166		
405_no_up	VAPEVLRVAEHRHKKGLMYPYIFHVLTKGEIKI	AVSIEDEANKOLPPAAL
gi 3005744	VAPEVLRVAEHRHKKGLMYPYIFHVLTKGEIKI	AVSIEDEANKOLPPAAL
405_hum	VAPEVLRVAEHRHKKGLMYPYIFHVLTKGEIKI	AVSIEDEANKOLPPAAL
gi 1136426 gnl PID d1012166		*****

## FIG. 11D

405_no_up	LYRPRQYVYGVLSLAESRKKTERLAFRKNRLPPEFSPLIIKEWAAKYG
gi 3005744	LYRPRQYVYGVLSLAESRKKTERLAFRKNRLPPEFSPLIIKEWAAKYG
405_hum	LYRPRQYVYGVLSLAESRKKTERLAFRKNRLPPEFSPLIIKEWAAKYG
gi 1136426 gnl PID d1012166	LYRPRQYVYGVLSLAESRKKTERLAFRKNRLPPEFSPLIIKEWAAKYG
405_no_up	KSPQTPELVEALAFREWTCPNLKRLWLKGAVEDKNRRMRAFLACMRSDTP
gi 3005744	KSPQTPELVEALAFREWTCPNLKRLWLKGAVEDKNRRMRAFLACMRSDTP
405_hum	KSPQTPELVEALAFREWTCPNLKRLWLKGAVEDKNRRMRAFLACMRSDTP
gi 1136426 gnl PID d1012166	KSPQTPELVEALAFREWTCPNLKRLWLKGAVEDKNRRMRAFLACMRSDTP
405_no_up	AMLNPANVPTHLMVLCVLRVYVQWPGARILRRQELDAFLAQALS PKLYE
gi 3005744	AMLNPANVPTHLMVLCVLRVYVQWPGARILRRQELDAFLAQALS PKLYE
405_hum	AMLNPANVPTHLMVLCVLRVYVQWPGARILRRQELDAFLAQALS PKLYE
gi 1136426 gnl PID d1012166	AMLNPANVPTHLMVLCVLRVYVQWPGARILRRQELDAFLAQALS PKLYE
405_no_up	PDQLQELKIENLDPRGIQLSALFMSGVDMALFANDACGQPIPWEHCPCPM
gi 3005744	PDQLQELKIENLDPRGIQLSALFMSGVDMALFANDACGQPIPWEHCPCPM
405_hum	PDQLQELKIENLDPRGIQLSALFMSGVDMALFANDACGQPIPWEHCPCPM
gi 1136426 gnl PID d1012166	PDQLQELKIENLDPRGIQLSALFMSGVDMALFANDACGQPIPWEHCPCPM

FIG. 11E

YFDGKLFOSKLLKASREKTPLIDICDQADQAANKVRQMSVLEGLFSR  
YFDGKLFOSKLLKASREKTPLIDICDQADQAANKVRQMSVLEGLFSR  
YFDGKLFOSKLLKASREKTPLIDICDQADQAANKVRQMSVLEGLFSR  
YFDGKLFOSKLLKASREKTPLIDICDQADQAANKVRQMSVLEGLFSR

```
405 no_up
gi|3005744
405 hum
gi|1136426|gnl|PID|d1012166
```

QNHLPFPFPFPALFYPASVYPRHFCVFPVSGCRGRFAGVCGFGGHYGE  
QSHTLFPFPFPALFYPASAYPRHFCVFPVSGCRGRFA-----  
QSHTLFPFPFPALFYPASAYPRHFCVFPVSGCRGRFAGVCGFGGHYGE  
QSHTLFPFPFPALFYPASAYPRHFCVFPVSGCRGRFAGVCGFGGHYGE  
QSHTLFPFPFPALFYPASAYPRHFCVFPVSGCRGRFAGVCGFGGHYGE  
\*\*\*\*\*  
\* \* \* \* \*  
\* \* \* \* \*

```
405 no up
gi|300574
405 hum
gi|1136426|gnl|PID|d1012166
```

TTVATGTPYRFRVTAASGHCCAFSGSDSNTSKSQGGVQVPIPSOGGKLEIA  
 -----GVQVPIPSOGGKLEIA  
 TTVATGTPYRFRVTAASGHCCAFSGSDSNTSKSQGGVQVPIPSOGGKLEIA  
 -----GVQVPIPSOGGKLEIA  
 TTVATGTPYRFRVTAASGHCCAFSGSDSNTSKSQGGVQVPIPSOGGKLEIA  
 -----GVQVPIPSOGGKLEIA  
 TTVATGTPYRFRVTAASGHCCAFSGSDSNTSKSQGGVQVPIPSOGGKLEIA  
 -----GVQVPIPSOGGKLEIA

```
405 no_up
gi|3005744
405 hum
gi|1136426|gnl|PID|d1012166
```

[illegible]

```
405 no up  
gi|300574  
405 hum  
gi|1136426|gnl|PID|d1012166
```

## FIG.11F

405\_no\_up  
gi|3005744  
405\_hum  
gi|1136426|gnl|PID|d1012166

405\_no\_up  
gi|3005744  
405\_hum  
gi|1136426|gnl|PID|d1012166

405\_no\_up  
gi|3005744  
405\_hum  
gi|1136426|gnl|PID|d1012166

GRYVGRGYKSGAIOGRPPYAASAEVAVAKELKSGESKSAVSSEGSIA  
GRYVGRGYKNOAAIOGRPPYAASAEVAVAKELKSGESKSSAMSSDGSIA  
GRYVGRGYKNOAAIOGRPPYAASAEVAVAKELKSGESKSSAMSSDGSIA  
GRYVGRGYKNOAAIOGRPPYAASAEVAVAKELKSGESKSSAMSSDGSIA  
\*\*\*\*\*.\*.\*\*\*\*\*  
ENGVMABEKEVPQLNGSTADTRAPSHSESALNNDSTCNTNPHLNALSTD  
ENGVMABEKEKPAQNMNGSTGDARAPSHSESALNNDSTCNTNPHLNALSTD  
ENGVMABEKEKPAQNMNGSTGDARAPSHSESALNNDSTCNTNPHLNALSTD  
ENGVIQR-----GEAGSPDERHG-----  
\*\*\*\*\*  
...:\*. \*

SACRREAALEAAVLNKEE  
SACRREAALEAAVLNKEE  
SACRREAALEAAVLNKEE  
-----

128/183

274 Human

Probe:

Pool of probes  
representing a 8Kb  
of the 274 Human  
contig

Target (total RNA  
sources)

1. Rat Bone
2. Rat Testis
3. Human cell line  
NB4

*FIG. 12*

129/183

**FIG. 13A**

Length of 274 25.3.99: 10427 bp; Listed from: 1 to: 10427;  
 Translated from: 211 to: 9996 (Entire region);  
 Genetic Code used: Universal;

**Frame 1**

CGA CAA AGG CAG CCT GTG CCG CAC AGT TGA GGG GTG CCG TGA GGA GCT GCA GAA  
 9 18 27 36 45 54

TCA GGC CAA TTT CTC CTT CCG TCC TCT CGT GTT AGA CAT GCT CAA TTT CCT CAT GGA TGC  
 63 72 81 90 99 108

CAT TCA GAC CAA CTT TCA CAG GCT TCC GCC GTG GGG AGC AGC AGC CCG GCA CAG CAG GCC  
 123 132 141 150 159 168

**OPEN**

CTC AGT GAA CTG CAC ACA GTG GAC AAG GTT GTG CAG ATG ACA GAC CAG CTG ATG GTT CCC  
 183 192 201 210 219 228

130/183

## FIG. 13B

T L G S Q E G A F E N V R M N Y S G D Q  
 ACC TTA GGC TCA CAG GAA GGT GCC TTT GAG AAT GTT CGG ATG AAT TAC AGC GGA GAC CAG  
 243 252 261 270 279 288

G Q T I R Q L I S A H V L R R V A M C V  
 GGC CAG ACT ATA CGG CAG CTG ATC AGT GCC CAC GTG CTC AGG CGG GTG GCT ATG TGT GTG  
 303 312 321 330 339 348

L S S P H G R Q H L A V S H E K G K I  
 CTT TCT TCC CCC CAT GGG CGC CAA CAT TTG GCT GTC AGC CAC GAG AAG GGC AAG ATC  
 363 372 381 390 399 408

T V L Q L S A L L K Q A D S S K R K L T  
 ACG GTG CTG CAG CTC TCG GCG CTC CTG AAG CAA GCA GAT TCC AGT AAG AGG AAG TTG ACT  
 423 432 441 450 459 468

L T R L A S A P V P F T V L S L T G N P  
 CTA ACC CGC TTG GCT TCT GCT CCA GTT CCC TTC ACC GTG CTG AGT CTT ACT GGA AAC CCC  
 483 492 501 510 519 528

C K E D Y L X V C G L K K D C H V L T F S  
 TGC AAG GAG GAC TAC CTG SCA GTG TGT GGG CTG CAT GTG CTC ACC TTC AGT  
 543 552 561 570 579 588  
  
 S S G S V S D H L V L H P Q L A T G N F  
 AGC TCG GGC TCT GTC TCC GAT CAC TTA GTG TTA 630  
 603 621 648  
  
 I I K A V W L P G S Q T E L A I V T A D  
 ATC ATC AAA GCT GTG TGG CTC CCT GGC TCT CAG ACC GAG CTA GCC ATT GTT GTC AAC GCA  
 663 672 681 690 708  
  
 F V K I Y D L S V D A L S P T F Y F L L  
 TTC GTC AAG ATT TAC GAC CTG TCT GTT GAT GCC TTG AGC CCC ACC TTC TAC TTC CTC CTG  
 723 732 741 759 768  
  
 P S S K I R D V T F L F N E E G K N I I  
 CCA AGC TCA AAG ATA AGG GAT GTC ACC TTC CTT TTC AAC GAG GAA GGC AAG AAC ATC ATT  
 783 792 801 810 819 828

FIG. 13C



## FIG. 13D

V I M S S A G Y T Q L M E E A S S A  
 GTC ATC ATG TCC TCT GGT GGG TAC ATG TAC ACA CAG CTC ATG GAG GAG GCC AGC AGT GCC  
 843 852 861 870 879 888

Q Q G P F Y V T N V L E I N H E D L K D  
 CAG CAG GGG CCC TTC TAT GTC ACT AAT GTG CTG GAA ATA AAT CAC GAG GAT CTG AAG GAC  
 903 912 921 930 939 948

S N S Q V A G G GGT GGC GTG TCT GTG TAC TAT TCA CAC GTG TTG CAG ATG  
 AGT AAC AGC CAG GTG GCA GGT 963 972 981 990 999 1008

L F F S Y S Q G CAG GGC AAG TCC TTT GCA GCC ACC GTC AGC AGG AGC ACT CTG  
 CTT TTC TTC AGC TAC AGT CAG 1023 1032 1041 1050 1059 1068

E V L Q L F P I N I K S S N G G S K T S  
 GAG GTG CTG CAG CTC TTC CCC ATC AAC ATC AAA AGC TCC AAT GGT GGC AGT AAG ACT TCT  
 1083 1092 1101 1110 1119 1128

133/183

## FIG. 13E

P A L C Q W S E V M N H P G L V C C V Q  
 CCC GCC CTT TGC CAG TGG TCT TGT GAG GTG ATG AAC CAC CCT GGC TTG GTG TGT TGT GTT CAG  
 1143 1152 1161 1170 1179 1188

Q T T G V P L V V M V K P D T F L I Q .E  
 CAA ACT ACT GGT GTG CCT CTG GTA GTC ATG GTG AAA CCA GAC ACT TTC CTC ATC CAG GAG  
 1203 1212 1221 1230 1239 1248

I K T L P A K A K I Q D M V A I R H T A  
 ATT AAG ACT CTT CCC GCC AAA GCA AAG ATC CAG GAC ATG GTT GCT ATT AGA CAC ACT GCC  
 1263 1272 1281 1290 1299 1308

C N E Q Q R T T M I L L C E D G S L R I  
 TGT AAT GAG CAG CGC ACC ACC ATG ATC CTG CTG TGT GAG GAT GGC AGC CTG CGA ATT  
 1323 1332 1341 1350 1359 1368

Y M A N V E N T S Y W L Q P S L Q P S S  
 TAT ATG GCC AAC GTG GAG AAC ACC TCT TAC TGG CTC CAG CCG TCT CTG CAG CCC AGC AGC  
 1383 1392 1401 1419 1428

## FIG. 13F

V I S I M K P V R K R K T A T I T A R T  
 GTC ATC AGC ATC ATG AAG CCT GTG CGA AAG CGC AAA ACA GCT ACA ATC ACA GCC CGC ACA  
 1443 1452 1461 1470 1479 1488

S S Q V T F P I D F F E H N Q Q L T D V  
 TCC AGC CAG GTG ACC TTC CCC ATT GAC TTC TTT GAA CAC AAC CAG CAG CTA ACG GAT GTG  
 1503 1512 1521 1530 1539 1548

E F G G N D L L Q V Y N A Q Q I K H R L  
 GAG TTT GGT GGT AAT GAC CTC CTG CAA GTG TAC AAT GCG CAA CAG ATA AAG CAC AGG CTC  
 1563 1572 1581 1590 1599 1608

N S T G M Y V A N T K P G G F T M E I S  
 AAC TCC ACT GGC ATG ATG TAT GTG GCG AAC ACC AAG CCT GGA GGC TTC ACC ATG GAG ATC AGT  
 1623 1632 1641 1650 1659 1668

N N S S T M V M T G M R I Q I G T Q A I  
 AAC AAC AGT AGC ACC ATG GTG ATG ACG GGC ATG CGG ATC CAG ATT GGC ACA CAG GCA ATC  
 1683 1692 1701 1710 1719 1728

## FIG. 13G

E R A P S Y I E I F G R T M Q L N L S R	
GAG CGA GCA CCG TCC TAC ATC GAG ATC TTT GGC AGA ACC ATG CAG CTT AAC CTG AGC CGC	1743 1752 1761 1770 1779 1788
S R W F D F P F T R E E A L Q A D R K L	
TCC CGC TGG TTT GAC TTC CCC TTC ACC AGA GAG GAA GCC CTG CAG GCT GAC CGG AAG CTG	1803 1812 1821 1830 1839 1848
N L F I G A S V D P A G V T M I D A V K	
AAC CTC TTC ATC GGT GCC TCT GTG GAT CCA GCT GGC GTC ACC ATG ATA GAT GAT GTA AAA	1863 1872 1881 1890 1899 1908
I Y G K K T K E Q F G W P D E P P E D F P	
ATT TAT GGC AAG ACT AAA GAG CAG TTT GGC TGG CCT GAT GAG CCC CCA GAA GAC TTC CCT	1923 1932 1941 1950 1959 1968
S A S V S S V C P P N L N Q S N G T G D	
TCT GCC TCT GTT AGC AGC GTC TGC CCT CCT AAC CTG AAC CAG AGC AAC GGC ACT GGA GAC	1983 1992 2001 2010 2019 2028

## FIG. 13H

S D S A A P A T S G T V L E R L V V S  
 AGT GAC TCA GCT GCT CCA GCT ACA ACC AGT GGC ACT GTC CTG GAG AGG CTG GTT GTG AGT  
 2043 2052 2061 2070 2079 2088  
  
 S L E A L E S C F A V G P I I E K E R N  
 TCT TTG GAA GCC CTG GAA AGC TGC TTT GCT GTC GGC CCA ATC ATT GAG AAG GAG AGA AAC  
 2103 2112 2121 2130 2139 2148  
  
 K H A A Q E L A T L L L S L P A P A S V  
 AAG CAC GCA GCC CAG GAG CTG GCC ACT TTG TTG CTC TCC CTG CCA GCG CCT GCC AGC GTC  
 2163 2172 2181 2190 2199 2208  
  
 Q Q Q S K S L L A S L H S S R S A Y H S  
 CAA CAG CAA TCA AAG AGC CTC CTG GCC AGC CTG CAC AGC AGT GGC TCG GCC TAC CAC AGC  
 2223 2232 2241 2250 2259 2268  
  
 H K D Q A L L S K A V Q C L N T S S K E  
 CAC AAG GAC CAG GCC TTG TTG AGC AAA GCT GTG TGT CTC AAC ACT TCC AGC AAA GAA  
 2283 2292 2301 2310 2319 2328

## FIG. 131

G K D L D P E V F Q R L V I T A R S I A	
GGC AAG GAC ITG GAC CCC GAG GTG TTC CAG CGT CTA GTA ATC ACA GCT CGC TCT ATT GCC	2388
	2379
	2370
	2361
	2352
	2403
V T R P N N L V H F T E S K L P Q M E T	
GTC ACA CGT CCC AAC AAC CTT GTG CAC TTT ACG GAG TCC AAG CTG CCC CAG ATG GAA ACA	2448
	2412
	2421
	2472
	2463
E G A E E G K E P Q K Q V E G D G C S F	
GAA GGA GCG GAG GAG GGG AAA GAG CCG CAG AAG CAG GTG GAA GGA GAC GGC TGT AGT TTC	2508
	2481
	2490
	2499
	2541
I T Q L V N H F W K L H A S K P K N A F	
ATC ACT CAG CTT GTA AAC CAC TTC TGG AAA CTC CAT GCA TCT AAG CCC AAG AAT GCC TTC	2568
	2532
	2550
	2559
	2619
	2628
L A P A C L P G L T H I E A T V N A L V	
CTG GCA CCT GCC TGC CTG CCA GGC CTT ACT CAT ATT GAA GCT ACG GTT AAT GCG CTG GTA	2619
	2583
	2601
	2619
	2628

## FIG. 13J

D I I H G Y C T C E L D C I N T A S K I  
 GAC ATT ATC CAT GGC TAT TGT ACC TGC GAG CTG GAC TGT ATC AAC ACA GCA TCC AAG ATC  
 2643 2652 2661 2670 2679 2688

Y M Q M L L C P D P A V S F S C K Q A L  
 TAC ATG CAG ATG CTG CTG TGT CCT GAC CCT GCT GTG AGC TTC TCC TGT AAA CAA GCT CTA  
 2703 2712 2721 2730 2739 2748

I R V L R P R N K R R H V T L P S S P R  
 ATT CGA GTC CTA AGG CCC AGG AAC AAG CGG AGA CAC GTG ACA TTG CCC TCC TCC CCC CGA  
 2763 2772 2781 2790 2799 2808

S N T P M G D K D D D D D A D E K M  
 AGC AAC ACT CCA ATG GGA GAC AAG GAT GAC GAT GAT GAT GAT GCG GAT GAG AAA ATG  
 2823 2832 2841 2850 2859 2868

Q S S G I P D G H I R Q E S Q E Q S E  
 CAG TCA TCG GGC ATC CCT GAT GGT GGT CAC ATC CGT CAG GAA AGC CAG GAA CAG AGT GAG  
 2883 2892 2901 2910 2919 2928

## FIG. 13K

V D H G D F E M V S E S M V L E T A E N  
 GTG GAC CAT GGA GAT TTT GAG ATG GTG TCT GAG TCG ATG GTC CTG GAA ACA GCT GAA AAT  
 2943 2952 2961 2970 2979 2988  
  
 V N N G N P S P R K A L L A G A E G F P  
 GTC AAC AAT GGC AAC CCC TCT CCC CGG AAA GCC CTG CTG GCC GGT GCT GAG GGC TTT CCT  
 3003 3012 3021 3030 3039 3048  
  
 P M L D I P P D A D D E T M V E L A I A  
 CCC ATG CTG GAC ATC CCA CCT GAT GCA GAC GAC ACC ATG GTT GAA CTA GCC ATT GCC  
 3063 3072 3081 3090 3099 3108  
  
 L S L Q Q D Q Q G S S S S A L G L Q S L  
 CTC AGC CTG CAG CAG GAC CAA GGC AGC AGC AGT GCC CTG GGC CTT CAG AGC CTG  
 3123 3132 3141 3150 3159 3168  
  
 G L S G Q A P S S S S L D A G T L S D T  
 GGA CTG TCG GGC CAG GCA CCC AGC TCT TCC TCT CTG GAC GCA GGA ACC CTC TCT GAC ACC  
 3183 3192 3201 3210 3219 3228



## FIG. 13L

T A S A P A S D D E G S T A A T D G S T  
ACA GCA TCA GCT CCA GCC TCA GAT GAC GAG GGC AGC ACT GCA GCA ACT GAT GAT GGC TCC ACC  
3243 3252 3261 3270 3279 3288

L R T S P A D H G G S V G S E S G G S A  
CTG CGG ACC TCG CCA GCC GAC CAT GGC GGT AGT GTG GGC TCA GAG AGC GGC GGA AGT GCA  
3303 3312 3321 3330 3339 3348

V D S V A G E H S V S G R S S A Y G D A  
GTG GAC TCG GTG GCT GGC GAG CAC AGT GTG TCT GGC CGG AGC AGT GCA TAT GGT GAC GCC  
3363 3372 3381 3390 3399 3408

T A E G H P A G P G S V S S S T G A I S  
ACA GCT GAG GGC CAC CCG GCT GGA CCA GGC AGC GTC AGC TCT AGC ACA GGC GCC ATC AGC  
3423 3432 3441 3450 3459 3468

T T T G H Q E G D G S E G E G E A E  
ACC ACT ACT GGC CAC CAG GAG GGA GAT GGA TCT GAG GGA GAA GAA GGC GAA GCT GAA  
3483 3492 3501 3510 3519 3528

## FIG. 13M

G D V H T S N R L H M V R L M L L E R L  
GGA GAT GTG CAC ACT AGC AAC AGG CTA CAC ATG GTT CGT CTA ATG CTG TTG GAA AGA TTA  
3543 3552 3561 3570 3579 3588

L Q T L P Q L R N V G G V R A I P Y M Q  
CTG CAG ACA CTG CCC CAG TTA CGG AAT GTC GGA GGT GTC CGG GCC ATC CCA TAC ATG CAG  
3603 3612 3621 3630 3639 3648

V I L M L T T D L D G E D E K D K G A L  
GTC ATT CTC ATG CTC ACT ACA GAT CTG GAT GGA GAA GAT GAG AAA GAC AAG GGA GCC CTG  
3663 3672 3681 3690 3699 3708

D N L L A Q L I A E L G M D K K D V S K  
GAC AAC CTG CTT ACA CAG CTC ATT GCT GAG CTG GGA ATG GAC AAA AAG GAT GTC TTC AAG  
3723 3732 3741 3750 3759 3768

K N E R S A L N E V H L V V M R L L S V  
AAG AAT GAA CGC AGT GCA TTG AAC GAA GTC CAT TTG GTC GTG ATG AGA CTC CTG AGT GTC  
3783 3792 3801 3810 3819 3828

## FIG. 13N

F M S R T K S G S K S S I C E S S S L I  
 TTC ATG TCC CGC ACC AAG TCT GGA TCC AAA TCT TCC ATC TGT GAG TCG TCT TCC CTC ATC  
 3843 3852 3861 3870 3879 3888

S S A T A A A L S S G A V D Y C L H V  
 TCC AGT GCC ACG GCA GCC CTG CTG AGC TCC GGC GCC GTG GAC TAT TGC CTC CAT GTG  
 3903 3912 3921 3930 3939 3948

L K S L L E Y W K G Q Q S E E E P V T T  
 CTC AAG TCT CTG CTG GAA TAC TGG AAG GGC CAG CAG AGT GAG GAG GAG CCT GTG ACC ACC  
 3963 3972 3981 3990 3999 4008

S Q L L K P H T S S S P P D M S P F F L  
 AGC CAA TTG CTG AAA CCA CAC ACC ACC TCA TCT CCA CCA GAT ATG AGC CCG TTC TTT CTC  
 4023 4032 4041 4050 4059 4068

R Q Y V K G H A A D V F E A Y T Q L L T  
 CGA CAA TAT GTG AAG GGT CAT GCT GCT GAT GTG TTC GAA GCC TAT ACC CAG CTG CTC ACA  
 4083 4092 4101 4110 4119 4128

## FIG. 130

E	M	V	L	R	L	P	Y	Q	I	K	K	I	A	D	T	S	S	R	I
GAA	ATG	GTA	CTG	AGG	CTT	CCC	TAC	CAG	ATC	AAG	AAG	ATC	GCA	GAT	ACC	AGC	TCT	CGA	ATC
		4143			4152			4161			4170			4179				4188	
P	P	P	V	F	D	H	S	W	F	Y	F	L	S	E	Y	L	M	I	Q
CCT	CCT	CCT	GTC	TTT	GAT	CAC	TCC	TGG	TTT	TAC	TTT	CTC	TCT	GAG	TAC	CTA	ATG	ATC	CAG
		4203			4212			4221			4230			4239				4248	
Q	T	P	F	V	R	R	Q	V	R	K	L	L	L	F	I	C	G	S	K
CAG	ACC	CCT	TTT	GTC	CGC	CGC	CAA	GTC	CGG	AAA	CTT	CTG	CTC	TTC	ATC	TGT	GGA	TCA	AAG
		4263			4272			4281			4290			4299				4308	
E	K	Y	R	Q	L	R	D	L	H	T	L	D	S	H	V	R	G	I	K
GAG	AAA	TAC	CGT	CAG	CTC	CGA	GAC	CTG	CAC	ACC	CTG	GAC	TCC	CAT	GTG	CGT	GGG	ATC	AAG
		4323			4332			4341			4350			4359				4368	
K	L	L	E	E	Q	G	I	F	L	R	A	S	V	V	T	A	S	S	G
AAG	CTG	CTG	GAG	GAA	CAA	GGC	ATC	TTT	CTC	AGG	GCC	AGC	AGC	GTG	GTT	ACA	GCC	AGT	TCC
		4383			4392			4401			4410			4419				4428	

## FIG. 13P

S	A	L	Q	Y	D	T	L	I	S	L	M	E	H	L	K	A	C	A	E
TCC	GCT	TIG	CAG	TAT	GAC	ACG	CTC	ATC	AGC	CTG	ATG	GAG	CAC	CTG	AAG	GCC	TGT	GCA	GAG
4443				4452				4461		4470				4479				4488	
I	A	A	Q	R	T	I	N	W	Q	K	F	C	I	K	D	D	S	V	L
ATT	GCC	GCC	CAG	CGC	ACC	ATC	AAC	TGG	CAG	AAG	TTC	TGC	ATC	AAA	GAT	GAC	TCT	GTC	CTG
4503				4512				4521		4530				4539				4548	
Y	F	L	L	Q	V	S	F	L	V	D	E	G	V	S	P	V	L	L	Q
TAC	TTC	CTC	CTG	CAA	GTG	AGT	TTC	CTG	GTA	GAT	GAG	GGG	GTG	TCC	CCT	GTG	CTA	CTG	CAG
4563				4572				4581		4590				4599				4608	
L	L	S	C	A	L	C	G	S	K	V	L	A	A	L	A	A	S	T	G
CTG	CTT	TCC	TGT	GCT	CTT	TGC	GGC	AGC	AAA	GTG	CTC	GCA	GCC	CTG	GCG	GCC	TCC	ACG	GGC
4623				4632				4641		4650				4659				4668	
S	S	S	V	A	S	S	A	P	P	A	A	S	S	G	Q	T	T	T	Q
TCC	TCC	AGT	GTG	GCT	TCC	TCA	GCC	CCT	CCT	GCA	GCC	AGT	TCT	GGA	CAA	ACC	ACA	ACT	CAG
4683				4692				4701		4710				4719				4728	

## FIG. 13Q

S K S S T K K S K K E E K E K E G E  
 TCC AAA TCC TCC ACT AAA AAG AGC AAG AAA GAA GAG AAA GAG AAA GGT GAG  
 4743 4752 4761 4770 4779 4788  
  
 S S G S Q E D Q L C T A L V N Q L N R F  
 AGC TCG GGT AGC CAG GAA GAC CAG CTG TGT ACA GCT CTC GTG AAC CAG CTG AAC AGA TTT  
 4803 4812 4821 4830 4839 4848  
  
 A D K E T L I Q F L R C F L L E S N S S  
 GCA GAC AAG GAG ACT CTG ATC CAG TTT CTG CGC TGC TTC CTA TTG GAG TCC AAC TCT TCG  
 4863 4872 4881 4890 4899 4908  
  
 S V R W Q A H C L T L H I Y R N S N K A  
 TCA GTA CGC TGG CAG GCC CAC CAC TGC CTG ACA CTG CAC ATC TAC AGA AAC TCC AAC AAG GCT  
 4923 4932 4941 4950 4959 4968  
  
 Q Q E L L L D L M W S I W P E L P A Y G  
 CAG CAG GAG CTC TTG CTA GAT CTC ATG TGG TCA ATC TGG CCA GAA CTC CCA GCT TAT CGT  
 4983 4992 5001 5010 5019 5028

## FIG. 13R

R K A A Q F V D L L G Y F S L K T A Q T  
 CGG AAG GGT GCC CAG TTT GTG GAT CTC CTG GGC TAT TTC TCC CTG AAA ACT GCA CAG ACA  
 5043 5052 5061 5070 5079 5088

E K K L K E Y S Q K A V E I L R T Q N H  
 GAG AAG AAG TTG AAG GAG TAC TCA CAG AAG GCC GTG GAG ATT CTG AGG ACT CAG AAC CAC  
 5103 5112 5121 5130 5139 5148

I L T N H P N S N I Y N T L S G L V E F  
 ATT CTC ACC AAT CAT CCC AAC TCC AAC ATT TAC AAT ACC TTG TCT GGC TTA GTG GAG TTT  
 5163 5172 5181 5190 5199 5208

D G Y Y L E S D P C L V C N N P E V P F  
 GAT GGC TAC TAC CTA GAG AGT GAT CCC TGC CTG GTG TGT AAT AAC CCT GAA GTG CCG TTT  
 5223 5232 5241 5250 5259 5268

C Y I K L S S I K V D T R Y T T Q Q V  
 TGT TAT ATC AAG CTG TCT TCC ATT AAA GTG GAC ACA CGG TAC ACC ACC CAG CAG GTG  
 5283 5292 5301 5310 5319 5328

## FIG. 135

V K L I G S H T I S K V T V K I G D L K  
 GTA AAG CTC ATT GGT AGT CAC ACC ATC AGC AAG GTG ACA GTG AAA ATT GGG GAC CTG AAG  
 5343 5352 5361 5370 5379 5388  
  
 R T K M V R T I N L Y Y N N R T V Q A I  
 CGG ACC AAG ATG GTG CGG ACA ATC AAC CTG TAT TAC AAC CAG ACT GTG CAG GCC ATT  
 5403 5412 5421 5430 5439 5448  
  
 V E L K N K P A R W H K A K K V Q L T P  
 GTG GAG TTG AAA AAT AAG CCA GCT CGC TGG CAC AAA GCC AAG AAA GTT CAG CTG ACG CCT  
 5463 5472 5481 5490 5499 5508  
  
 G Q T E V K I D L P L P I V A S N L M I  
 GGA CAG ACA GAA GTC AAG ATT GAC CTG CCT CTG CCC ATT GTG GCC TCT AAC CTG ATG ATC  
 5523 5532 5541 5550 5559 5568  
  
 E F A D F Y E N Y Q A S T E T L Q C P R  
 GAG TTT GCA GAC TTC TAT GAG AAC TAC CAG GCT TCC ACA GAG ACC CTT CAG TGT CCT CGC  
 5583 5592 5601 5610 5619 5628



## FIG. 13T

C S A S V P A N P G V C G N C G E N V Y	TGC AGT GCC TCT GTC CCC GCC AAT CCQ GGC GTC TGT GGC AAC TGT GGC AAC TAT GTA TAC	5643	5652	5661	5670	5679	5688
Q C H K C R S I N Y D E K D P F L C N A	CAG TGT CAC AAG TGC AGG TCC ATC AAC TAT GAT GAG AAG GAT CCC TTC CTC TGC AAT GCT	5703	5712	5721	5730	5739	5748
C G F C K Y A R F D F M L Y A K P C C P	TGT GGC TTT TGT AAA TAT GCC CGC TTT GAC TTC ATG CTT TAT GCC AAA CCT TGC TGC CCA	5763	5772	5781	5790	5799	5808
V D P I E N E E D R K K A V S N I N T L	GTG GAT CCC ATC GAG AAC GAG GAA GAT CGG AAG AAG GCC GTT TCC AAC ATC AAC ACA CTT	5823	5832	5841	5850	5859	5868
L D K A D R V Y H Q L M G H R P Q L E N	CTG GAC AAA GCT GAC CGC GTG TAC CAT CAG CTC ATG GGA CAC CGG CCC CAG CTG GAG AAC	5883	5892	5901	5910	5919	5928

## FIG. 13U

L L C K V N E A P E K P Q E D S G T A  
CTG CTC TGC AAA GTG AAT GAA GCA GCT CCA GAA AAG CCA CAG GAC TCG GGA ACA GCG

G G I S S T T A S V N R Y I L Q L A Q E  
GGA GGC ATC AGC TCC ACT TCA GCC AGT GTG AAT GCG TAC ATC CTA CAG CTG GCG CAG GAG  
6003 6002 6012 6021 6020 6030 6039 6048

Y C G D C K N S F D E L S K I I Q K V F  
TAT TGT GGA GAC TGT AAC AAC TCA TTT GAC GAG CTC TCC AAA ATC CAG AAA GTC TTC  
6063 6072 6081 6090 6099 6108

A S R K E L L E Y D L Q Q R E A A T K S  
GCT TCG CGC AAA GAG CTG TTA GAG TAT GAC CTG CAG AGA GAA GCG GCC ACC AAG TCA

S R T S V Q P T F T A S Q Y R A L S V L  
TTC CGG ACA TCC GTG CAG CCC AGC TTT ACT GCC AGC CAG TAC GGT GCG TTT GTC TCT TTA

## FIG. 13V

G C G H T S S T K C Y G C A S A V T E H	
GGC TGT GGC CAC ACC TCC TCC ACC AAG TGC TAT GGC TGT GCC TCA GCT GTC ACA GAG CAT	6243 6252 6261 6270 6279 6288
C I T L L R A L A T N P A L R H I L V S	
TGT ATC ACA CTG CTA AGG GCC CTG GCC ACC AAT CCA GCC CTG AGA CAC ATC CTC GTC TTC	6303 6312 6321 6330 6339 6348
Q G L I R E L F D Y N L R R G A A A I R	
CAG GGC CTC ATC CGA GAG CTC TTT GAT TAC AAC CTC CGC AGA GGC GCG GCA GCC ATC CGC	6363 6372 6381 6390 6399 6408
E E V R Q L M C L L T R D N P E A T Q Q	
GAG GAG GTC CGC CAG CTC ATG TGC CTC CTG ACT AGA GAC AAC CCA GAG GCC ACC CAG CAA	6423 6432 6441 6450 6459 6468
M N D L I I G K V S T A L K G H W A N P	
ATG AAT GAC CTG ATC ATT GGC AAA GTC TCC ACT GCA CTG AAG GGC CAC TGG GCT AAT CCT	6483 6492 6501 6510 6519 6528

## FIG. 13W

D L A S S L Q Y E M L L L L T D S I S K E  
 GAC CTG GCT AGC AGC CTG TAT GAG ATG CTG ACA GAC TCC ATC TTC AAG GAG  
 6543 6552 6561 6570 6579 6588

D S C W E L R L R C A L S L F L M A V N  
 GAC AGC TGC TGG GAA CTC CGC TTA CGC TGT GCA CTC AGC CTT TTC CTC ATG GCT GTG AAC  
 6603 6612 6621 6630 6639 6648

I K T P V V V E N I T L M C L R I L Q K  
 ATC AAA ACG CCA GTG GTT GTT GAG AAC ATC ACC CTC ATG TGT CTG CGG ATC TTA CAG AAG  
 6663 6672 6681 6690 6699 6708

L I K P P A P T S K K N K D V P V E A L  
 CTG ATT AAA CCA CCT GCC CCA ACC AGC AAC AAG AAC AAG GAC GTC CCT GTG GAG GCC CTC  
 6723 6732 6741 6750 6759 6768

T T V K P Y C N E I H A Q A Q L W L K R  
 ACC ACG GTG AAG CCG TAC TGC AAT GAG ATC CAC GCG CAG GCT CAG CTG TGG CTC AAG CGA  
 6783 6792 6801 6810 6819 6828

## FIG. 13X

D P K A S Y E A W K K C L P I R G V D G  
 GAT CCT AAG GCA TCC TAC GAA GCC TGG AAG TGC CTG CCT ATC CGA GCG GTA GAT GGC  
 6843 6852 6861 6870 6879 6888

N G K S P S K S E L H R L Y L T E K Y V  
 AAC GGG AAA TCC CCC AGC AAG TCC GAG CTC CAC CGG CTC TAC TTG ACT GAG AAG TAT GTG  
 6903 6912 6921 6930 6939 6948

W R W K Q F L S R R G K R T T P L D L K  
 TGG AGG TGG AAA CAG TTC CTG AGT CGG CGG AAG AGG ACC ACC CCA CTT GAC CTC AAG  
 6963 6972 6981 6990 6999 7008

L G H N W L R Q V L F T P A T Q A A R  
 CTG GGC CAC A'C AAC TGG CTG CGG CAG GTA CTC TTT ACC CCG GCA ACA CAG GCA GCA CGG  
 7023 7032 7041 7050 7059 7068

Q A A C T I V E A L A T V P S R K Q Q V  
 CAG GCA GCC TGT ACT ATT GTG GAA GCT CTT GCT ACT CTC CCC AGC CGC AAG CAG CAG GTG  
 7083 7092 7101 7110 7119 7128

## FIG. 13Y

L D L L L T S Y L L D E L S V A G E C A A E  
 CTT GAC CTC ACC AGT TAC CTG GAC GAG CTG AGT GTG GCT GGT GCA GAG  
 7143 7152 7161 7170 7179 7188

Y L A L L Y Q K L I A S C H W K V Y L A A  
 TAC CTG GCT CTC TAC CAG AAG CTC ATC GGC TCC TGC CAC TGG AAA GTC TAC CTG GCT GCT  
 7203 7212 7221 7230 7239 7248

R G V L L P Y V G N L I T K E I A R L L A  
 CGG GGA GTC CTG CCC TAC GTG GGC AAC CTC ATC ACC AAG GAA ATC GCC CGC CTG CTG GCC  
 7263 7272 7281 7290 7299 7308

L E E A T L S T D L Q Q G Y A L K S L T  
 TTG GAG GAG GCC ACA CTG AGC ACA GAC CTG CAG CAG GGA TAT GCC CTC AAG AGT CTC ACA  
 7323 7332 7341 7350 7359 7368

G L L S S F V E V E S I K R H F K S R L  
 GGC CTC CTC TCC TCC TTT GTA GAG GTG GAG TCC ATC AAG CGT CAT TTC AAG AGC CGC TTG  
 7383 7392 7401 7410 7419 7428

## FIG.13Z

V G T V L N G Y L C L R K L V L Q R T K  
 GTG GGC ACT GTG CTG AAT GGG TAC TGC TTG CGG AAA CTA GTG CTG CAG AGG ACC AAG  
 7443 7452 7461 7470 7479 7488

L I D E T Q D M L L E M L E D M T T G T  
 CTC ATC GAT GAG ACC CAG GAC ATG CTG GAG ATG CTG GAG GAC ATG ACC ACA GGT ACA  
 7503 7512 7521 7530 7539 7548

E S E T K A A F M A V C I E T A K R Y N L  
 GAG TCT GAA ACC AAG GCC TTC ATG GCT GTG TGC ATT GAG ACA GCC AAG CGC TAC AAT CTG  
 7563 7572 7581 7590 7599 7608

D D Y R T P V F I F E R L C S I I Y P E  
 GAT GAC TAC CGG ACT CCT GTG TTC ATC TTT GAG AGG CTG TGC AGC ATC ATC TAC CCT GAG  
 7623 7632 7641 7650 7659 7668

E N E V T E F F V T L E K D P Q Q E D F  
 GAA AAT GAA GTC ACT GAG TTC TTT GTG ACC CTG GAG AAG GAC CCC CAG CAA GAG GAC TTT  
 7683 7692 7701 7710 7719 7728

155/183

## FIG. 13A

L	Q	G	R	M	P	G	N	P	Y	S	S	N	E	P	G	I	G	P	L
CTA	CAG	GGC	AGG	ATG	CCC	GGG	AAC	CCA	TAT	AGC	AGC	AAT	GAA	CCA	GGC	ATT	GGG	CCT	CTT
	7743				7752			7761		7770				7779				7788	
M	R	D	I	K	N	K	I	C	Q	D	C	D	L	V	A	L	L	E	D
ATG	AGG	GAC	ATA	AAG	AAC	AAG	ATC	TGC	CAG	GAC	TGT	GAC	CTG	GTG	GCT	CTT	CTG	GAG	GAC
	7803				7812			7821		7830				7839				7848	
D	S	G	M	E	L	L	V	N	N	K	I	I	S	L	D	L	P	V	A
GAC	AGT	GGG	ATG	GAG	CTT	CTA	GTG	AAC	AAT	AAA	ATT	ATC	AGT	CTG	GAT	CTT	CCT	GTG	GCT
	7863				7872			7881		7890				7899				7908	
E	V	Y	K	K	V	W	C	T	T	N	E	G	E	P	M	R	I	V	Y
GAG	GTT	TAC	AAG	AAG	GTC	TGG	TGT	ACG	ACC	AAT	GAG	GGA	GAG	COC	ATG	AGG	ATT	GTT	TAT
	7923				7932			7941		7950				7959				7968	
R	M	R	G	L	L	G	D	A	T	E	E	F	I	E	S	L	D	S	T
CGA	ATG	CGG	GGG	CTA	CTG	GGC	GAT	GCC	ACC	GAG	GAG	TTT	ATC	GAG	TCC	CTG	GAC	TCC	ACC
	7983				7992			8001		8010				8019				8028	



## FIG. 13BB

T D E E E D E E E V Y R M A G V M A Q C  
 ACA GAT GAA GAA GAC GAA GAG GAA GTG TAC AGA ATG GCC GGT GTG ATG GCC CAG TGC  
 8043 8052 8061 8070 8079 8088

G G L Q C M L N R L A G V K D F K Q G R  
 GGG GGT CTG CAG TGC ATG CTG AAC AGA CTG GCA GGC ATG AAA GAT TTT AAG CAG GGG CGC  
 8103 8112 8121 8130 8139 8148

H L L T V L L K L F S Y C V K V N R  
 CAC CTT CTA ACA GTG CTC CTG AAG CTG TTC AGT TAC TGT GTG AAG GTG AAA CTC AAT CGG  
 8163 8172 8181 8190 8199 8208

Q Q L V K L E M N T L N V M L G T L N L  
 CAA CAG TTG GTC AAG CTG GAA ATG AAC ACC TTG AAT GTC ATG TTG GGG ACT TTA AAC TTG  
 8223 8232 8241 8250 8259 8268

A L V A E Q E S K D S G G A A V A E Q V  
 GCT CTG GTA GCT GAG CAA GAG AGC AAG GAC AGC GGA GGA GCA GCG GTG GCT GAG CAG GTG  
 8283 8292 8301 8310 8319 8328

## FIG. 13CC

L S I M E I I L D E S N A E P L S E D K	
CTG AGC ATC ATG GAG ATC ATT CTA GAT GAG TCC AAT GCA GAG CCC CTG AGT GAG GAC AAG	8388
	8379
	8370
	8361
	8352
	8343
G N L L L L T G D K D Q L V M L L D Q I N	
GGC AAC CTC CTC ACA GGT GAC AAG GAT CAA CTG CTG ATG CTC TTG GAC CAG ATC AAC	8448
	8439
	8430
	8421
	8412
	8403
S T F V R S N P S V L Q G L L R I I P Y	
AGC ACC TTC GTT CGT TCC AAT CCC AGC GTG CTG CAG GGT TTG CTG CGT ATC ATC CCA TAC	8508
	8499
	8490
	8481
	8472
	8463
L S F G E V E K M Q I L V E R F K P Y C	
CTG TCC TTC GGA GAG GTG GAG AAG ATG CAG ATC CTG GTG GAG CGG TTC AAG CCA TAC TGC	8568
	8559
	8550
	8541
	8532
	8523
S F D K Y D E D H S G D D K V F L D C F	
AGC TTC GAT AAG TAC GAT GAA GAC CAT AGT GGG GAT GAC AAA GTC TTC CTG GAT TGC TTC	8628
	8619
	8610
	8601
	8592
	8583

## FIG. 13DD

C K I A A G I K N S N S G H Q L K D L I  
 TGC AAG ATT GCT GGT ATC AAG AAC AAC AGC AAT GGT CAT CAG CTG AAG GAC CTC ATC  
 8643 8652 8661 8670 8679 8688

L Q K G I T Q S A L D Y M K K H I P S A  
 CTG CAG AAG GGA ATC ACC CAG AGT GCC CTG GAC TAC ATG AAA AAG CAT ATC CCC AGT GGC  
 8703 8712 8721 8730 8739 8748

K N L D A D I W K K F L S R P A L P F I  
 AAG AAT TTG GAT GCT GAC ATC TGG AAA AAG TTT TTG TCT CGG CCT GCT CTG CCA TTT ATT  
 8763 8772 8781 8790 8799 8808

L R L L R G L A M Q H P A T Q V L I G T  
 TTG AGG CTT CTT CGG GGT CTG GCC ATG CAA CAC CCT GCC ACC CAG GTT CTG ATT GGA ACG  
 8823 8832 8841 8850 8859 8868

D S I T S L H K L E Q V S S D E G I G T  
 GAC TCC ATC ACA AGC CTG CAT AAG CTA GAG CAG GTG TCC AGT GAT GAG GGC ATT GGG ACC  
 8883 8892 8901 8910 8919 8928

159/183

## FIG. 13EE

L A E N L L E A L R E H P D V N K K I D	
TTG GCC GAG AAC CTC CTG GAG GCC CTA AGA GAA CAC CCA GAT GTG AAC AAG AAG ATT GAC	8943 8952 8970 8979 8988
A A R R E T R A E K K R M A M A M R Q K	
GCA GCC CGC AGA GAG ACC CGG ATG GGC ATG AAG CGC ATG GCC ATG AGG CAG AAG	9003 9012 9021 9030 9039 9048
A L G T L G M T T N E K G Q V T K T A	
GCT CTG GGA ACT CTG GGC ATG ACG ACA AAC GAG AAG GGT CAG GTG GTG ACC AAG ACT GCG	9063 9072 9081 9090 9099 9108
L L K Q M E E L I E E P G L T C C I C R	
CTC CTG AAG CAA ATG GAA GAG CTG ATT GAG GAG CCC GGA CTC ACG TGC TGC ATC TGC AGG	9123 9132 9141 9150 9159 9168
E G G T K F Q P T K V L G I Y T F T K R	
GAA GGG GGT ACC AAG TTC CAG CCC ACG AAG GTC TTA GGC ATT TAC ACT TTC ACC AAG CCG	9183 9192 9201 9210 9219 9228

## FIG. 13FF

V A L E E M E N K P R K Q Q G Y S T V S  
 CTG GCC TTG GAG GAG ATG GAG AAC AAG CCT CGG AAA CAG CAG GGC TAC AGC ACT GTG TCC  
 9243 9252 9261 9270 9279 9288  
  
 H F N I V H Y D C H L A A V R L A R G R  
 CAT TTC AAC ATC GTG CAC TAT GAC TGT CAC CTG GCT GCT GTG AGG TTG GCT AGA GGC CGG  
 9303 9312 9321 9330 9339 9348  
  
 E E W E S A A L Q N A N T K C N G L L P  
 GAA GAG TGG GAA AGT GCT GCC CTG CAG AAC GCC AAC ACG GAG TGC AAT GGA CTC CTT CCG  
 9363 9372 9381 9390 9399 9408  
  
 V W G P H V P E S A F A T C L A R H N T  
 GTC TGG GGC CCC CAT GTC CCT GAG TCA GCT TTT GCC ACT TGT TTA GCT AGG CAC AAC ACT  
 9423 9432 9441 9450 9459 9468  
  
 Y L Q E C T G Q R E P T Y Q L N I H D I  
 TAC CTC CAG GAG TGT ACC GGC CAG CGG GAG CCC ACG TAC CAG CTC AAC ATC CAC GAC ATC  
 9483 9492 9501 9510 9519 9528

## FIG. 13GG

K L L F L R F A M E Q S F S A D T G G G  
 AAA CTG CTT TTC CTG CGC TTC GCC ATG GAG CAG TCA TTC AGT GCA GAC ACC GGT GGG GGC  
 9543 9552 9561 9570 9579 9588

G R E S N I H L I P Y I I H T V L Y V L  
 GGT CGG GAG AGC AAC ATC CAC CTG ATC CCG TAC ATC ATT CAC ACT GTG CTT TAC GTC CTG  
 9603 9612 9621 9630 9639 9648

N T T R A T S R E E K N L Q G F L E Q P  
 AAC ACG ACC CGA GCA ACG TCC CGG GAG GAG AAG AAC CTC CAA GGC TTC CTG GAG CAG CCC  
 9663 9672 9681 9690 9699 9708

R E K W T E S A F D V D G P H Y F T I L  
 AGA GAG AAG TGG ACA GAG AGT GCC TTT GAT GTA GAT GGG CCC CAC TAC TTC ACC ATC TTA  
 9723 9732 9741 9750 9759 9768

A L H V L P P E Q W K A T R V E I L R R  
 GCC CTG CAT GTC CTC CCC CCT GAG CAG TGG AAA GCC ACT CGA GTG GAG ATC CTA CGC AGG  
 9783 9792 9801 9810 9819 9828

## FIG. 13HH

L L V A S H A R A V A P G A T R L T D  
 CTG CTG GCC TCG CAT GCC CGG GCA GTG GCT CCA GGA GGA ACC AGG TTG ACA GAT  
 9843 9852 9861 9870 9879 9888

K A V K D Y S A Y R S S L L F W A L V D  
 AAG GCG GTG AAA GAC TAT TCT TCC TTG CTG TGG GCT CTC GTT GAT  
 9903 9912 9921 9930 9939 9948

L I Y N M F K K V P T S N T  
 CTC ATT TAT AAC ATG TTT AAG GTG CCC ACT AGT AAC ACT TGA GGG CGG GCT GGT CCT  
 9963 9972 9981 9990 9999 10008

NGC TCT CTG GCT GAG TAC ATT CCG CCA CAA ATG ACA TGC CCA TTC TAC GAA GCT GCC GAC  
 10023 10032 10041 10050 10059 10068

AAA GCC CCT GAA AAN CTT CCA GGA GGA GTT CAT GCC AGT GGA GAC CTT CTC AGA GTT CCT  
 10083 10092 10101 10110 10119 10128

*FIG. 13 II*

CGA CGC AGC AGG TCT TCT ATC AGA AAT CAC CGA CCC AGA GAG CTT CCT AAA GGA CCT GTT  
 10143 10152 10161 10170 10179 10188

GAA CTC AGT CC<sup>2</sup> CTG ACC ATC CCC ATG GAA ATG AGA TCG CGG TGA CGA GAT TGA AGC TAG  
 10203 10212 10221 10230 10239 10248

CTT GCA TTC TCC CCT CAT TCA TCT CTC CGT GTG CAC TCA TTT CCC CCA CGG ATG CTG CAT  
 10263 10272 10281 10290 10299 10308

TAG CAC CCC CTC CTC CCT CAG TTT TCT TGG AGT GGC TTG GGG TTT GTA GGC TTC CTG TTC  
 10323 10332 10341 10350 10359 10368

TAT CAC GCG CGC GTG TGT GTG TGT GTG TGT GTG TGT GTG TGT GTG TG  
 10383 10392 10401 10410 10419



## FIG. 14A

274: CIUSTAL X (1.64b) Multiple Sequence Alignment

```

274
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|
MSAHSGGTDMNSVVKALIILNRTGALNKNEVNVLLKAITRCEHDFEESN
-----

274
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|
FTQFYTAFAALADKIMQIKTICQTICQLHDATAVLIRFIYRLPRVSV
-----

274
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|
YETKWLGLKALMCEGRCFASASSMEDYNANVLKSCRPSTTKSI
-----

274
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|
MPSSSSSGSGASNDKESPKSEIKRSRDLSSVTLQQLIAPLEPGKMTWVP
-----

274
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|
LSERVTDCTEQILANVEYFQEQNGVDTLLDVCVSLPIILNRYRSKIMETI
-----

```

## FIG. 14B

274: CIUSTAL X (1.64b) Multiple Sequence Alignment

274	-----	-----
gi 3413886 dbj BAA32307	NGGKSLYPLTQVEATVKSMMHMLTDLTILSOAALIEHQPLTPSRIE	
gi 4426611 gb AAD20450	-----	
274	-----	-----
gi 3413886 dbj BAA32307	PLSMCGIAALYNVLTSTATSVLGHMSQASSQKQTASTSQSGVGSSGG	
gi 4426611 gb AAD20450	-----	
274	-----	-----
gi 3413886 dbj BAA32307	QSNKHDDFEDQACSIYNKALEIYSNIGHMFKTSARIHVYQNHLCYGSWL	
gi 4426611 gb AAD20450	-----	
274	-----	-----
gi 3413886 dbj BAA32307	LISGIQMGASCGSSDSASKASKATKSGSEAGTAPTTPIARVNLF	
gi 4426611 gb AAD20450	-----	
274	-----	-----
gi 3413886 dbj BAA32307	KVQQGFELMDAIAHHSTKLSELIEDLKVEACGQSLSETELEPPAQFD	
gi 4426611 gb AAD20450	-----	

## FIG. 14C

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
ILQNYSLERIVRLMTATLHOLFATVAYRKACTLKASAKORTECE  
-----

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
PISYSDSTTYFNDSLSCSDNSEDDSEYLGHWFKETLSPETHDNNANTS  
-----

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
TOERAEQKSALVPKLDEPHEYLDLADI FCFDQFLANRHAYMORYKAG  
-----

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
VSDQOMLLMANI IKDFDRVMRNETDQGSNAPAAASAGAGTSAGASTKWQ  
-----

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
TSMIRFSGAAGRYIHNLI STLSLLSEQLSNLLQHL SI SPWSTDNTWPLQ  
-----

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
VYFSTLSVIVQILLKPKTQEKERACLSVWHRLINTLVEGVCSNTASDSD  
-----

**FIG. 14D**

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
YEDLNIEHAQALLFLFHS LNMOKKSILLTTAGGVI RCAEVCRGISED RP

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
VKNSQIMLLSRILLFLEYLMKHLYNAPPELDOVRWNI FSVSSMPDTQKI

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
TDLNCRTKLNSYCRHDIEENFRKSAGEYSSIRPTFYSLVMGDPE

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
ISYTAQEFFKLDGLAWNFI LCTPDKLKYPLVLDALTDILSITDMSMY

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
SMHNLCAIQYCFYTIANKNLGLPPSTSHVESILKAERSPNLHSLMWS

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
IRLPASSHYLVVSSLIKQMYTQYAEITLWTHVGDIGADIKYSLKQ

FIG. 14E

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
-----  
-----  
NSQMGNGTFRLSDI.IIFDSLVAHQAVAWANKKGLKWPKESEDA  
-----

-----  
-----  
-----  
AGEQSAGTSLTSSNDPELYSSNESIDDKLKQDDGKGLSSDLQKY  
-----

-----  
-----  
-----  
VKLMNSYQLLSVIVRGQMLKQLSSSTFEKALNLIIVIVSDKPAIML  
-----

-----  
-----  
-----  
ELHAAFLKLLPNEKDQLIANEWPCLMVNDFAFNGKHQHPVEPYILN  
-----

-----  
-----  
-----  
LTRGSNYSTLHTLKHKLSILQLFELLPHRTANAEEVETQLKQLLI  
-----

-----  
-----  
-----  
SSMLDMRTDYLGQHSCHLREILSGLTQEAQKLLLYEHNVGVGYTRM  
-----

FIG. 14F

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

-----  
RQPQAAGPSCGAPLDQDRAMFNESMLFVLTMTMFKWLEKPVAVQAM  
-----

-----  
RQFFKQQRSGSILTTLLLSFTGTSIPVSYARKMLQFVNRLFQLSLQA  
-----

-----  
LVECFSELATVDVARLKQWLGHIIYGNVSTDVSTSEALDTTCRML  
-----

-----  
TSILQPSSSSSNAQPTTNMATVSAMPSISDQLDPMEIEYDCGTAA  
-----

-----  
TGGAANTSOILSIWQAQDPFSESSQAQCHSDSERNGALLISFV  
-----

-----  
KSLVKDQSKASQIAPPLFQALLQGLTILISPPQCCDFADVLQIMI  
-----

## FIG. 14G

274

gi | 3413886 | dbj | BAA32307 |  
 gi | 4426611 | gb | AAD20450 |

-----  
 RGHVALFNTLLWLELAKLQLPDKHLKHAENVSAQLRYLSELLQSI  
 -----

274

gi | 3413886 | dbj | BAA32307 |  
 gi | 4426611 | gb | AAD20450 |

-----  
 GFRGSRQHNPFWDDELQTDIDELYDELAEEGEQDSLDDSDDEDILN  
 -----

274

gi | 3413886 | dbj | BAA32307 |  
 gi | 4426611 | gb | AAD20450 |

-----  
 ALLSKAVQCLNTSS-----KEGKOLDPEVFQRLVITARSLAVTRPNNLV  
 -----  
 VLLQFVNRELOSMQPKLEKLETLREIDPEAFYRLVLMVRGIANARPOSIA  
 -----

## FIG. 14H

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|HFTESKLPQMETEGAECKEPQKQVEGDCSFITQVNVHFWKLGASKPKN  
-----YDIVPTLMGIVLELHKVTPTL  
KICVENN-----

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|AFLAPACLPGLTHIENTVNAIVDI IHGYCTCELDICINFRSKI INQMLLCP  
-----DEPNIVKRGLCQPETIVHCLIVEIMYGFALADPGQVGRMTKYFIDLKHC  
-----

274

gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|DPAVSFSCQALIRVLRPNRKRHVTLFSSPRSNTPMGDKDDDDDDDADE  
-----ASVISHSAKEALILLSPRMRKRKVAIVTPACSTPTPTSTMTQALQAAA  
-----



**FIG. 4I**

```

274      RMQSSGIPDGCHIRQESQEQSEVDHGDGFEMVSESMVLETAENVNNGNPSF
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|

274      RKALLAGAEGFPFPMIDIPPADDETVELAIAISLQDQCGSSSSALGLQ
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|

274      LEAFMGG--GFPRLGLPEADDEAINDIAIAISLQHQHGDANALQSLQ
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|

274      SIG--LSGQAPSSSIDMGTLSDTTASAPASDDDEGSTAATDGTSTLRTSPAD
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|

274      GLANLQGIQATAMAAAVNAANVSLGSSDDDEGSNVATDGTSTLRTSPA
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|

274      HGGVSGSESGSANDSVAGEHSVSGRSSAYGDATAEHPAGPGSVSS--
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|

274      QPGFSGVAGSGGVAVLSAMSSSEANEADOKLS--KLHDIRITVLESIIQ
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|

274      -TG--AISTTTGHQEGDGEGEGEAGEAGDVHTSNRLHWRLMLLERLIQ
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|

274      TLPLQNRVGGVRAIPYQVILMLTTDLGDEKKGALDNLISQLIAELG
gi|3413886|dbj|BAA32307|
gi|4426611|gb|AAD20450|

```



274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

TTTTQWKLGSHTISKVTKIGDKRTKWRVTINLYNNRTVOAIVELK  
TTTTQWKLGSHTISKVTKIGDKRTKWRVTINLYNNRTVOAIVELK  
TTTTTKYKIVOCHTISKIYRIADLKRTRVTINLYNNRVSQVAIVELK  
\*\*\* \*\* \*\*\*\*\* \*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\*



## FIG. 14M

274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

RELEFDYMLRGGAAAREEVRQMLCLLTRDNPETQOQNDLLIIGKVSTALK  
RELEFDYMLRGGAAAREEVRQMLCLLTRDNPETQOQNDLLIIGKVSTALK  
SELAHNLRRGTGLIOEEVRLVVLTKONAEACMHLLQVLTTRVKNALM  
\*\*::\*\*\*\*\*::\*\*\*\*\*::\*\*\*\*\*::\*\*\*\*\*::\*\*\*\*\*::\*\*\*\*\*  
GHWANPDASSLOYEMLLTDSISKEDSCWELRLCALSLFLMNVNIKTP  
GHWANPDASSLOYEMLLTDSISKEDSCWELRLCALSLFLMNVNIKTP  
GSIPLISLEAAVHQEMTLEVLGGDDICWEYKLVIFELFISNCRIPR-  
\*::\*:\*\*\*::\*\* \*\* \*\* \*\* \*\* \*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\* \*\*  
VVENITLMCLIRILQKLIKPPAPT SKKNKDVPEALTTPKPYCNEIHAQA  
VVENITLMCLIRILQKLIKPPAPT SKKNKDVPEALTTPKPYCNEIHAQA  
GPVTAVLHPCLIRIMQNLICPVLP GSKENOKVATTDLCSMKMEGNT-VDY  
\*::\*\*\*\*\*::\*\* \*\* \*\* \*\* \*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\*  
QLWLKRDPKASYEAWKKCLPIRG-----VDGNGKSPSKS--E  
QLWLKRDPKASYEAWKKCLPIRG-----IDGNGKAPSKS--E  
RAWLNSDRNHEYAAWSKMPSSNQAKLKNAKQONVAAGSGSDAPPKSRRE  
\*::\*:\*\*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\*  
LHRLXLYTEKYVVRWQKFLSRGKRTTFLDLKLGHNWNLQVLTFTATQAA  
LHRLXLYTEKYVVRWQKFLSRGKRTTFLDLKLGHNWNLQVLTFTATQAA  
VRVAFTEKYGKRWREVLDDQKRVIKPLVFN-----KWIQPLLFNANSRFG  
::\*\*\*\*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\*  
ROACTIVELATVP SRKQOVLDDLTYSYDELSVAGECAAEYALYQKLI  
ROACTIVELATVP SRKQOVLDDLTYSYDELSVAGECAAEYALYQKLI  
ROLACSLSSISRTNRRQOALNMLTSLFKHVGEAGASAEYMLYKNA  
\*\* \*\* \*\* \*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\* \*\*::\*\* \*\* \*\*

gi|3413886|abJ|BAC122301

gi | 3413666 | ab | AAD20450 |

gi | 3413000 | ad | 20450 |

g1|34I3000|uBJ|BIB22

gr | 3413686 | uuj : BHA:22-  
...11486611 = b | 33320450 |

gi|3413886|dbj|BAA2321.1|1-142cc11|ab|AAD204501

TSAHWKVYLARGVLPVGNLTITKEIARLLAGEAIISSA...

TGLSSFVEVESIKRHFKSRLVGIVLNGILECERCEV

TQKEFMNQHWYHCHTCNMINTVGVCVGRZ

GAKEGSCQALSRRIGSNEVRDSAGTGS III STAGE

VLTRKDSL TNERIAVLTKLLEPYREILQHQDQNEE

-----DEPARTMENT OF JUSTICE-----

## FIG. 140

274	gi 3413886 dbj BAA32307	NVRMNSGDQGTINQLISAHLRVRVAMCVLSSPHGRROHLAVSHEKGKI
	gi 4426611 gb AAD20450	-----
		NVRMNSGDQGTINKHLLSGVRRVAFCLSSPHGRROQLAVSHEKGKV
274	gi 3413886 dbj BAA32307	TVLQLSALLKQADSRRKRLTLRLASAPVPFTVLSLTGNPCKEDYLXVCG
	gi 4426611 gb AAD20450	-----
		TVLQLSALLKQADASKRKLTLTOLSSAPIACTVLSLAANPCNEDCLAVCG
274	gi 3413886 dbj BAA32307	LKCHVLTFFSSGSVDHLVHPOLATGNFIIKAVWLFGSQTELAVITAD
	gi 4426611 gb AAD20450	-----
		LKECHILTFSSGSSTNEHIVVNPOLENGNYIKKAVLPGSQTLILAVVTSD
274	gi 3413886 dbj BAA32307	FVKIYDLSVDALSPTFYLLPSSKIRDVTFLENEEGKNIIVIMSSAGYNY
	gi 4426611 gb AAD20450	-----
		YVKIYDLAVDTYSPKYYLAVAGKIRDCFTFYQDGNYNMLTFASS-GYIY
274	gi 3413886 dbj BAA32307	TQIMEEASSAQGPFVTVNVLEINHEDLKDSNSQVAGGVSYYVSHVLQM
	gi 4426611 gb AAD20450	-----
		TQQLDQOSLAVHGDFVTVNTLELSHOIKDIAGHIGGGVSIYYSHTLQL
274	gi 3413886 dbj BAA32307	LFESYSGKSEFAATVRSRTLEVLQFLPFIINTKSSNGSKTSP--ALCQWSE
	gi 4426611 gb AAD20450	-----
		LFYSYSGRSFFSPLTNVSEGVKGIYHLDTNSASKSASKRPLQPLVQWTE

## FIG. 14P

274  
 VMNHPGLVCCVQTTGVLVNVKPEDTELQEIKTLPKAKIODVAIRH  
 GI|3413886|dbj|BAA32307|  
 GI|4426611|gb|AAD20450|

274  
 VTGHPGLVYASMTSNPILMITPERIYLQEIKAOSAKSRIMDVUGIRH  
 GI|3413886|dbj|BAA32307|  
 GI|4426611|gb|AAD20450|

274  
 TACNEQORTMILLCEGDSLRIYMANVENTSYMLQPSLOPSS---VISI  
 GI|3413886|dbj|BAA32307|  
 GI|4426611|gb|AAD20450|

274  
 AVAG-VEKTTILLCEGDSLRIYSAQPEYTSFWLSPQVQFGNQLYSSTL  
 GI|3413886|dbj|BAA32307|  
 GI|4426611|gb|AAD20450|

274  
 MKP-----VRKRKTATITARTSS-----QVTFPIDFEEHN  
 GI|3413886|dbj|BAA32307|  
 GI|4426611|gb|AAD20450|

274  
 MAKGGSGSTS KSKSNTASGKMTSRKASQOQKPTAGGPVFPIDFEEHC  
 GI|3413886|dbj|BAA32307|  
 GI|4426611|gb|AAD20450|

274  
 OQLTDVEFGNDLLQVYNAQIKHRLNSTGMVYVANTKPGGFTMEISNNS  
 GI|3413886|dbj|BAA32307|  
 GI|4426611|gb|AAD20450|

274  
 NMLADVEFGNDLLQIYNKQKLTRELFSTGMFVASTRSNGFTLEVINNDP  
 GI|3413886|dbj|BAA32307|  
 GI|4426611|gb|AAD20450|

274  
 TMVMTGWRIOIGTOAIERAPSYIEIFGRTMQLNLSRWFDFPFTREAL  
 GI|3413886|dbj|BAA32307|  
 GI|4426611|gb|AAD20450|

274  
 NVVIVGIRVLIGTQDVQAPQSVTILGRTTPTFVRARWEDIPLTREML  
 GI|3413886|dbj|BAA32307|  
 GI|4426611|gb|AAD20450|

274  
 QADRKMLFTIGASVDPAGVTMIDAVKIYKTKEQFCGWPDEPPEDFESASV  
 GI|3413886|dbj|BAA32307|  
 GI|4426611|gb|AAD20450|

274  
 QSDKLLKVFAKAPDEPHVTILLDCIEVYGKSKELVGWPEDESD---VTV  
 GI|3413886|dbj|BAA32307|  
 GI|4426611|gb|AAD20450|



**SUBSTITUTE SHEET (RULE 26)**

274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

```
QVLI GDSITS LHKLEQVSDEGI GTLAENLLEALREHPDVNKKIDAARR
QVLIGDTSIPNIHKL EQVSDEGI GTLAENLLEALREHPDVNKKIDAARR
QVAISKDIPII HRL EQVSDEHVGSIAENLLEALSTSDTATAARVQVQRD
***** :*:***** . . . :. *
```

## FIG. 14S

274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

274  
gi|3413886|dbj|BAA32307|  
gi|4426611|gb|AAD20450|

ETRAEKKRMAVRQKALGTGLTTRNEKGQVVTALLKQMEELIEEPL  
ETRAEKKRMAVRQKALGTGLTTRNEKGQVVTALLKQMEELIEEPL  
ETRAEKKRMAVRQKALGTGLTTRNEKGQVVTAKGILQIEKRDETGL  
ETRAEKKRLAMATREKQALDGLGRTNEKGQVAKGILQIEKRDETGL  
\*\*\*\*\*:\*\*\* \*\* \*::\*\*\*\* \*\*\*\*\*: \* :\*\*\*\*\*: \* :\*\*\*  
\*\*\*\*\*:\*\*\* \*\* \*::\*\*\*\* \*\*\*\*\*: \* :\*\*\*\*\*: \* :\*\*\*

TCCICREGGKTQPTKVLGIYFTTKRVALEEMENKPKQOGYSTVSHFN  
TCCICREGGKTQPTKVLGIYFTTKRVALEEMENKPKQOGYSTVSHFN  
TCCICREGGKTQPTKVLGIYFTTKRVALEEMENKPKQOGYSTVSHFN  
TCFICREG-YACQPDKVLGIYFTTKRVALEEMENKPKQOGYSTVSHFN  
\*\*\*\*\* \*\* \*\*\*\*\*:\*\*\* \*\* \*::\*\*\*\* \*\*\*\*\*: \* :\*\*\*\*\*: \* :\*\*\*

VHDXCHLAANRLARGREWEESALONANTKCNGLLPVWGPHVPEAFATC  
VHDXCHLAANRLARGREWEESALONANTKCNGLLPVWGPHVPEAFATC  
VHDXCHLAANRLARGREWEESALONANTKCNGLLPVWGPHVPEAFATC  
VHDXCHLAANRLARGREWEESALONANTKCNGLLPVWGPHVPEAFATC  
\*\*\*\*\* \*\* \*\*\*\*\*:\*\*\* \*\* \*::\*\*\*\* \*\*\*\*\*: \* :\*\*\*\*\*: \* :\*\*\*

LARHNTYLOECTGQREPTYQLNIHDIKLLFIRFAMEQSFADTGGGGRS  
LARHNTYLOECTGQREPTYQLNIHDIKLLFIRFAMEQSFADTGGGGRS  
LARHNTYLOECTGQREPTYQLNIHDIKLLFIRFAMEQSFADTGGGGRS  
LARHNTYLOECTGQREPTYQLNIHDIKLLFIRFAMEQSFADTGGGGRS  
\*\*\*\*\* \*\* \*\*\*\*\*:\*\*\* \*\* \*::\*\*\*\* \*\*\*\*\*: \* :\*\*\*\*\*: \* :\*\*\*

NIHLIPYIHTVLYVLTTRATSRREKNLOGLFQOP-REKWTESAFVDG  
NIHLIPYIHTVLYVLTTRATSRREKNLOGLFQOP-REKWTESAFVDG  
NIHLIPYIHTVLYVLTTRATSRREKNLOGLFQOP-REKWTESAFVDG  
NIHLIPYIHTVLYVLTTRATSRREKNLOGLFQOP-REKWTESAFVDG  
\*\*\*\*\* \*\* \*\*\*\*\*:\*\*\* \*\* \*::\*\*\*\* \*\*\*\*\*: \* :\*\*\*\*\*: \* :\*\*\*

## FIG. 14T

274	gi 3413886 dbj BAA32307	PHYFTIALHVLPPQWKATVEIRRLIVASHARAVAPGGATR---LT
	gi 4426611 gb AAD20450	PYFTVIALHILPPQWEATVEIRRLIVTSOARAVAPGGATR---LT
		PLFNATISLSHRELWNKHLKHLKMTAVAGGRHUSPAVLCKALLAFA
		* ::: : * * * : : * : : : : * * * : : * * * : :
274	gi 3413886 dbj BAA32307	DKAVKDYSAVRSSLLFWALVDLIYN-MFKKVPSTNT---GGMWSCSLAEYIRH
	gi 4426611 gb AAD20450	DRQVKDYTKFPLMWWALVDLIYDNLFTVSTPKED-WFISLFDYIRK
		* : * * * : * : : * * * * * : : * * * : :
274	gi 3413886 dbj BAA32307	NDMPIYEAADKALKTFQEEFNPVETSEFLDVAGLLSEITDPDSFLKDLL
	gi 4426611 gb AAD20450	NDEALLKSTDSILQTLTEEFPLCTSFVEEFDVAGLLHLIEHPDNFIEEIL
274	gi 3413886 dbj BAA32307	-----
	gi 4426611 gb AAD20450	NSVP-----
		AALFSTSSSN

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/11066

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C 12Q 1/68

US CL : 435/6

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 435/7.1, 514/44, 435/69.1, 536/23.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

US PAT, Biosis, embase, caplus, medline, lifesci.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y,P	US 5,882,925 A (FALB) 16 March 1999 see entire document.	1-12
Y,P	US 5,861,249 A (BEACH et al) 19 January 1999 see entire.	5, 9-12
Y	US 5,445, 941 A ( YANG) 29 August 1995 see entire document.	1-12
Y	US 5,599,708 A (MUNDY et al) 04 February 1997 see entire document.	1-12
Y,P	US 5,763,416 A (BONADIO et al) 09 June 1998 see entire document.	1-12

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"R" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is considered with one or more other such documents, such combination being obvious to a person skilled in the art

"A\*" document member of the same patent family

Date of the actual completion of the international search

24 SEPTEMBER 1999

Date of mailing of the international search report

29 OCT 1999

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

ANNE B. BARNES

Telephone No. (703) 308-0254

Form PCT/ISA/210 (second sheet/July 1992)\*

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/11066

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/11066

## BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-12 and 43, drawn to method or process for identifying genes whose expression is responsive to a specific cue or cues (claims 1-12) and a method for advancing research in or studies of bone development (claim 43).  
Group II, claim(s) 16-19, drawn to method for determining risk of developing osteoporosis or low or high bone density or other factors causing or contributing to osteoporosis or lower levels of osteoblasts and chondrocytes or other conditions involving mechanical stress or lack thereof.

Group III, claim(s) 19, drawn to method for testing a medicament for or gene therapy approach to osteoporosis.

Group IV, claim(s) 20, 36-40, 44 and 45, drawn to method of treating, preventing or controlling osteoporosis.

Group V, claim(s) 22, drawn to a cellular-based model of osteoporosis or mechanical stress, or lack thereof.

Group VI, claim(s) 23-28, drawn to nucleic acid, probes, primers (claims 23-25 and 28) and to a vector (claims 26 and 27).

Group VII, claim(s) 29-34 and 41-42, drawn to expression product (polypeptide), composition comprising same, and related method of making same.

Group VIII, claim(s) 34 and 35, drawn to antibody and related composition.

Claim 21 is considered to be generic to the inventions of groups VI, VII, and VIII. Accordingly, it will be searched with any or all of the Groups depending upon the payment of additional fees for having searches conducted for each of said groups.

Election of Species

In the event that applicant elects to pay additional fees for the search of inventions corresponding to Groups III, IV, VI, VII, and VIII, the following election of species will apply.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

Claim 21: a) gene or portion thereof;  
b) protein or portion thereof; and  
c) antibody or portion thereof.

Group IV: a) protein 274; and  
b) protein 405.

Group VI: a) nucleic acid that encodes human protein 608;  
b) nucleic acid that encodes human protein 405;  
c) nucleic acid that encodes human protein 274; and  
d) probe or primer that specifically hybridizes to  
i) a);  
ii) b); and  
iii) c).

Group VII: a) protein 274;  
b) protein 405; and  
c) protein 608.

Group VIII: a) antibody that is elicited by protein 274;  
b) antibody elicited by protein 405; and  
c) antibody elicited by protein 608.

The inventions listed as Groups I-VIII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The inventions are directed to different methods and compounds. The methods comprised of different methods steps, requiring different reagents and resulting in different end products. The compounds are not related by a special technical feature, e.g., the probe/primer need not encode any amino acid sequence nor antibody and can also hybridize specifically to other

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/11066

sequences with equal affinity.

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons: The species are drawn to different proteins, nucleic acids, probes/primers, and antibodies which each have different physical composition and activities.

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.